ANGLIA RUSKIN UNIVERSITY

THE GENETIC AND MOLECULAR BASIS OF MELANISM IN THE GREY SQUIRREL *(SCIURUS CAROLINENSIS)*

HELEN REBECCA MCROBIE

A thesis in partial fulfilment of the requirements of Anglia Ruskin University for the degree of Doctor of Philosophy

Submitted: September 2014

Acknowledgments

I am very grateful to Anglia Ruskin University for giving me the opportunity to undertake this project. Thank you goes to my first supervisor, Dr Peter Coussons, for support and advice throughout. Furthermore, this work would not have been possible without the cell culture laboratory which was established and managed by Peter. Thank you also goes to my second supervisor, Dr Linda King, for invaluable encouragement and guidance throughout. I am also grateful for the support and advice from my advisers, Dr Cristina Fanutti, Dr Martyn Symmons and Dr John O'Brien. Thank you also to Dr Alison Thomas for initiating the project. I am grateful for the assistance of Mark D'Arcy, Angela Wheatley, Dr Nancy Moncrief, Luca Lapini, Dr Caray Walker, Dr Nick Pugh, Dr PhilWarburton and Dr Phillip Pugh.

An enormous thank you goes to my family for their inspiration, support and patience throughout this project. In particular I would like to say thank you to my dear husband Allan McRobie. Al's constant encouragement and kindness have made all the difference. Thank you.

ANGLIA RUSKIN UNIVERSITY

ABSTRACT

FACULTY OF SCIENCE AND TECHNOLOGY

DOCTOR OF PHILOSOPHY

THE GENETIC AND MOLECULAR BASIS OF MELANISM IN THE GREY SQUIRREL (*SCIURUS CAROLINENSIS*)

HELEN REBECCA MCROBIE

SEPTEMBER 2014

The grey squirrel (*Sciurus carolinensis*) has wildtype and melanic (dark) colour morphs. Melanism is associated with variations in the melanocortin-1 receptor (*MC1R*) gene in a number of species. The MC1R protein is a G-protein coupled receptor, predominantly expressed in melanocytes, where it is a key regulator of pigment production. To investigate the genetic and molecular basis of melanism, the *MC1R* genes of the wildtype and melanic grey squirrel were sequenced. The wildtype (MC1R-wt) and melanic (MC1RΔ24) variants of the MC1R were then functionally characterised in a cell-based assay.

The *MC1R* gene of the grey squirrel was found to have a 24 base pair (bp) deletion associated with melanism. The MC1R is typically activated by its agonist, the alpha-melanocyte stimulating hormone (α-MSH), which stimulates dark pigment production by raising intracellular cAMP levels. Conversely, the MC1R is inactivated by its inverse agonist, the agouti signalling protein (ASIP), which stops dark pigment production by lowering intracellular cAMP levels. To investigate the effects that the 24 bp deletion have on receptor function, *MC1R-wt* and *MC1RΔ24* genes were transfected into HEK293 cells. Cells expressing either MC1R-wt or MC1RΔ24 were stimulated with α-MSH or ASIP and intracellular cAMP levels were measured. Unstimulated MC1RΔ24 cells showed higher basal activity than the MC1R-wt cells. Both MC1R-wt and MC1RΔ24 cells responded to α-MSH with a concentration-dependent increase in intracellular cAMP. However, while the MC1R-wt cells responded to ASIP with a concentration-dependent *decrease* in intracellular cAMP, MC1RΔ24 cells responded with an *increase* in cAMP.

Melanism in the grey squirrel is associated with a 24 bp deletion in the *MC1R*. Cells expressing MC1RΔ24 have higher basal levels of cAMP than MC1R-wt cells. ASIP acts as an inverse agonist to the MC1R-wt but as an agonist to the MC1RΔ24. As MC1RΔ24 cells have higher levels of cAMP, and higher levels of cAMP lead to dark pigment production, the 24 bp deletion is the likely molecular cause of melanism in the grey squirrel.

Key words: *Sciurus carolinensis*, grey squirrel, melanism, melanocortin-1 receptor.

Publications

McRobie, H.R., Thomas, A. and Kelly, J., 2009. The genetic basis of melanism in the gray squirrel (*Sciurus carolinensis*). *The Journal of Heredity,* 100 (6), pp.709-714.

McRobie, H.R., King, L.M., Fanutti, C., Coussons, P.J., Moncrief, N.D. and Thomas, A.P., 2014. Melanocortin-1 receptor (MC1R) gene sequence variation and melanism in the gray (*Sciurus carolinensis*), fox (*Sciurus niger*), and red (*Sciurus vulgaris*) squirrel. *The Journal of Heredity,* 105 (3), pp.423-428.

McRobie, H.R., King, L.M., Fanutti, C., Symmons, M.F. and Coussons, P.J., 2014. Agouti signalling protein is an inverse agonist to the wildtype and agonist to the melanic variant of the melanocortin-1 receptor in the grey squirrel (*Sciurus carolinensis*). *FEBS Letters.588 (14) pp.2335-2343.*

Contents

Copyright Declaration

The Genetic and Molecular Basis of Melanism in the Grey Squirrel (*Sciurus carolinensis*)

by

Helen Rebecca McRobie

Attention is drawn to the fact that copyright of this thesis rests with Anglia Ruskin University for one year and thereafter with Helen Rebecca McRobie. This copy of the thesis has been supplied on condition that anyone who consults it is bound by copyright.

List of Figures

List of Tables

Abbreviations

µl	microlitre
µM	micromolar
Å	ångstrom unit
A2AR	adenosine A_{2a} receptor
ACTH	adrenocorticotropin
AgRP	agouti related protein
AP1	activator protein-1
AP2	activator protein-2
ASIP	agouti signalling protein
ATP	adenosine triphosphate
BLAST	basic local alignment search tool
BLASTN	basic local alignment tool search nucleotide
bp	base pairs
BSA	bovine serum albumin
cAMP	cyclic adenosine monophosphate
CASP	Critical Assessment of Structure Prediction
CBP	CREB-binding protein
CREB	cAMP response element binding protein
C-terminal	carboxyl terminal
DCT	dopachrome tautomerase
ddH_2O	double distilled water
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
E boxes	enhancer boxes
EC_{50}	concentration giving half the maximal response
ECL	extracellular loop

EDTA	ethylene diamine tetra acetic acid
ELISA	Enzyme-Linked Immunosorbent Assay
ER	endoplasmic reticulum
g	gram
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GDP	guanosine diphosphate
GFP	green fluorescent protein
GPCR	G-protein coupled receptor
G-protein	guanosine nucleotide binding protein
GRK	GPCR kinase
GTP	guanosine triphosphate
HEK293T	human embryonic kidney 293T cells
ICL	intracellular loop
kb	kilo base
kDa	kilo dalton
KIT	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog
L	litre
LB	Luria-Bertani
L-DOPA	L-3, 4-dihydroxyphenylalanine
MC1R	melanocortin-1 receptor
MC2R	melanocortin-2 receptor
MC3R	melanocortin-3 receptor
MC4R	melanocortin-4 receptor
MC5R	melanocortin-5 receptor
MCR	melanocortin receptor
mg	milligram
$MgCl_2$	magnesium chloride
MITF	microphthalmia transcription factor

ml	millilitre
mm	millimetre
mM	millimolar
mRNA	messenger ribonucleic acid
NCBI	National Center for Biotechnology Information
ng	nanogram
nm	nanometre
NMR	nuclear magnetic resonance
N-terminal	amino terminal
^{o}C	degrees Celsius
PBS	phosphate buffered saline
PDB	Protein Data Bank
Pitx1	paired-like homeodomain 1
PKA	protein kinase A
POMC	pro-opiomelanocortin
PSI-BLAST	position specific iterated basic local alignment search tool
RMSD	root mean square deviation
rpm	revolutions per minute
SCOP	Structural Classification Of Proteins
SDS-PAGE	sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SNP	single nucleotide polymorphism
SP1	specificity protein-1
TBE	tris borate EDTA
TBS	tris buffered saline
TM	transmembrane helix
TRP1	tyrosinase related protein
TYR	tyrosinase
UV	ultraviolet

V	volts
wt	wildtype
α-MSH	alpha melanocyte stimulating hormone

Amino acids

Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic acid	Glu	E
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Chapter One

General Introduction

"The colours of many animals seem adapted to their purposes of concealing themselves…The final cause of these colours is easily understood, as they serve some purposes of the animal, but the efficient cause would seem almost beyond conjecture" (Darwin, 1796). These are the words Erasmus Darwin wrote in "Zoonomia; or, the Laws of Organic Life" in reference to the colours found in the animal kingdom. The term "final cause" refers to the biological purpose and the term "efficient cause" refers to the biological mechanism. Thanks to the astonishing progress in the field of Biology, biological mechanisms are now no longer "beyond conjecture" and this thesis is an attempt to understand the "efficient cause" of melanism (darkening) in the grey squirrel (*Sciurus carolinensis*).

1.1 Thesis Overview

This thesis primarily concerns one particular protein called the melanocortin-1 receptor (MC1R) which is a key regulator of pigmentation. The MC1R lies in the membrane of pigment-producing cells where it functions as a molecular switch. Variations in the gene encoding this protein have been associated with colour variations in many vertebrate species including humans (Lin and Fisher, 2007). The present study investigates genetic variations in the *MC1R* gene of the grey squirrel and goes on to consider the effect these variations have on the protein's ability to function as a switch. Chapter one is a general literature review which begins with a brief consideration of the squirrel family and goes on to consider two key cases of melanism. This is followed by a brief overview of the cell biology of melanin formation with a consideration of the genetic basis of pigmentation variations, particularly those involving the *MC1R* gene. The introduction goes on to consider the molecular mechanisms by which the MC1R protein functions as a switch and how computer modelling

can be used to investigate such molecules. Chapter two describes the materials and methods used in the study. Chapter three presents results of investigations into the genetics of melanism in the grey squirrel. Chapter four presents results of a secondary investigation into the genetics of melanism in two other species, the fox (*Sciurus niger*) and red (*Sciurus vulgaris*) squirrels. The three dimensional structure of the MC1R of the grey squirrel is investigated in Chapter five using computer modelling predictions. Chapter six presents results of *in vitro* experiments investigating the function of the MC1R of the grey squirrel using a cell based system and Chapter seven is a general discussion of all the findings in the study.

1.2 The Squirrel Family

Squirrels are conspicuous and charismatic creatures and can be found on all continents apart from Antarctica and Australasia (Gurnell and Gurnell, 1987). The squirrel family, Sciuridae, is a diverse family of mammals with more than 270 species belonging to 50 genera (Herron, Castoe and Parkinson, 2004). Squirrels can be divided into three broad categories; tree squirrels, ground-dwelling squirrels and flying squirrels (Gurnell and Gurnell, 1987). It is thought that all three types of squirrel evolved from a common ancestor which lived in North America in the Eocene epoch between 54 and 37 million years ago (Black, 1995). Flying squirrels are thought to have diversified from the main lineage during the Eocene epoch and tree and ground-dwelling squirrels are thought to have diversified widely towards the end of the Oligocene epoch between 37 and 25 million years ago (Black, 1995). In this study, three species of tree squirrel are considered with the primary focus on the grey squirrel and with a secondary focus on both the fox squirrel (*Sciurus niger*) and the red squirrel (*Sciurus vulgaris*). All three of these tree squirrel species belong to the genus *Sciurus*, family Sciuridae, order Rodentia, class Mammalia, phylum Chordata and kingdom Animalia.

1.2.1 The grey, fox and red squirrel: phenotypes

Tree squirrels are diurnal and arboreal, living in cold temperate forests of the northern hemisphere where they typically feed on nuts, tree seeds, buds and berries (Gurnell and Gurnell, 1987). There is much variation in the body weights within these squirrel species according to the time of year and food availability. The grey squirrel has a head and body length between 380-525 mm, a tail length between 150-250 mm and a body weight between 300-710 g (Gurnell and Gurnell, 1987). The fox squirrel is larger than the grey squirrel with a

head and body length between 450-700 mm, a tail length between 200-330 mm and a body weight between 700-1230 g (Gurnell and Gurnell, 1987). The red squirrel is the smallest of the three species with a head and body length between 180-240 mm, tail length between 140-195 mm and a body weight between 250-340 g (Shorten, 1954). All three species produce two litters a year, one in December to February and one in May to June, typically producing three or four young per litter.

The wildtype grey squirrel, as its name suggests, has an overall grey appearance with a grizzled dorsum and russet regions present on the backs of the paws and also along the dorsal-ventral boundary, as shown in figures 1.1 A and 1.4. The ventrum is paler than the rest of the body, and is usually white but may be pale grey (Gurnell and Gurnell, 1987).

Figure 1.1. Colour morphs of the grey squirrel (*Sciurus carolinensis*)

A) Wildtype and B) melanic grey squirrel. Photographs by Trimming (2011) and Whippey (2012).

The wildtype fox squirrel is similar to the grey squirrel but has a more reddish colour overall (see figure 1.2 A).

Figure 1.2. Two colour morphs of the fox squirrel (*Sciurus niger*)

A) Wildtype and B) melanic fox squirrel. Photographs by Loser (2007).

Red squirrel colouration varies between bright red to brown to black with all morphs having a white ventrum (see figure 1.3).

Figure 1.3. Colour morphs of the red squirrel (*Sciurus vulgaris*)

A) Red, B) dark red, C) brown and D) melanic morphs of the red squirrel demonstrating the spectrum of colour morphs of this species. Photographs by Trimming (2011) and Whippey (2012).

The overall colour of these three species tends to match the tree bark of the dominant trees in their particular region; the grey and fox squirrels match the grey tree bark of the maple and oak (see figure 1.4), and the red squirrel matches the pine, spruce, fir and cedar of their habitat (Smith, 1981).

Figure 1.4. Grey squirrel (*Sciurus carolinensis*) near an oak tree

A grey squirrel sitting on an oak tree demonstrating the match between its coat colour and its habitat. Photograph by Jaga (2008).

This colour match suggests that squirrel colouration offers camouflage which may protect from predation. All three species of squirrel are hunted by a range of predators including hawks, weasels, foxes and snakes (Gurnell and Gurnell, 1987). All three species have melanic (darkened) variants (see figures 1.1 B, 1.2 B and 1.3 B, C, D). In the grey squirrel, the melanic variant is generally reported to be in the north of its range in North America, whereas melanic variants of the fox squirrel are found in the south of its range, also in North America (Gurnell and Gurnell, 1987). Melanism in the red squirrel is not as obviously defined, having a spectrum of colours ranging from red to black across its range in Eurasia (figure 1.3). Melanic variants of the red squirrel are particularly common in the Alps and Pyrenees (Gurnell and Gurnell, 1987).

1.2.2 Geographic distribution of squirrels

The grey squirrel is found in the eastern regions of North America, extending as far west as Kansas in the USA, and as far north as Saskatchewan in Canada. Introductions have also occurred in California, Montana, Oregon and Washington in the USA and Quebec, New Brunswick, British Columbia, Manitoba, Nova Scotia, Ontario and Saskatchewan in Canada (Barkalow and Shorten, 1973). Further introductions have occurred in other countries on other continents, including England, South Africa and Italy (Gurnell and Gurnell, 1987). The fox squirrel has a similar range to the grey squirrel but it extends further west, as far as Montana in the USA but not as far north as the grey squirrel in Saskatchewan in Canada (Gurnell and Gurnell, 1987). Both the grey and fox squirrel are found in dense, mature, hardwood forests and mixed forests, often bordering on prairie or farmland. The grey squirrel, in particular, is highly adaptable and can live in both urban and suburban areas (Gurnell and Gurnell, 1987). The red squirrel has a vast range across Eurasia, from the Mediterranean, the Caucasus Mountains, the southern Urals, the Altai Mountains of central Mongolia to north-east China. The red squirrel is found in dense, coniferous or broadleaved forests across its range (Gurnell and Gurnell, 1987).

The grey squirrel was introduced to the British Isles from North America where it has become a successful invasive species. The earliest reference to the presence of the grey squirrel in the British Isles is in 1830, but this early sighting is not verified (Middleton, 1931). It is generally accepted that the first verified sighting of a grey squirrel in the British Isles was in 1876 in Cheshire (Middleton, 1931). Grey squirrels were repeatedly introduced from both the USA and Canada between 1876 and 1910 and subsequently relocated within

the country. Although there is no definite documentation, it is reported that twelve melanic squirrels were introduced to Woburn some time after the introduction of wildtype grey squirrels in the area (Shorten, 1954). The first map to document the spread of the melanic grey squirrel in the British Isles was produced in 1954 and is shown in figure 1.5 (Shorten, 1954).

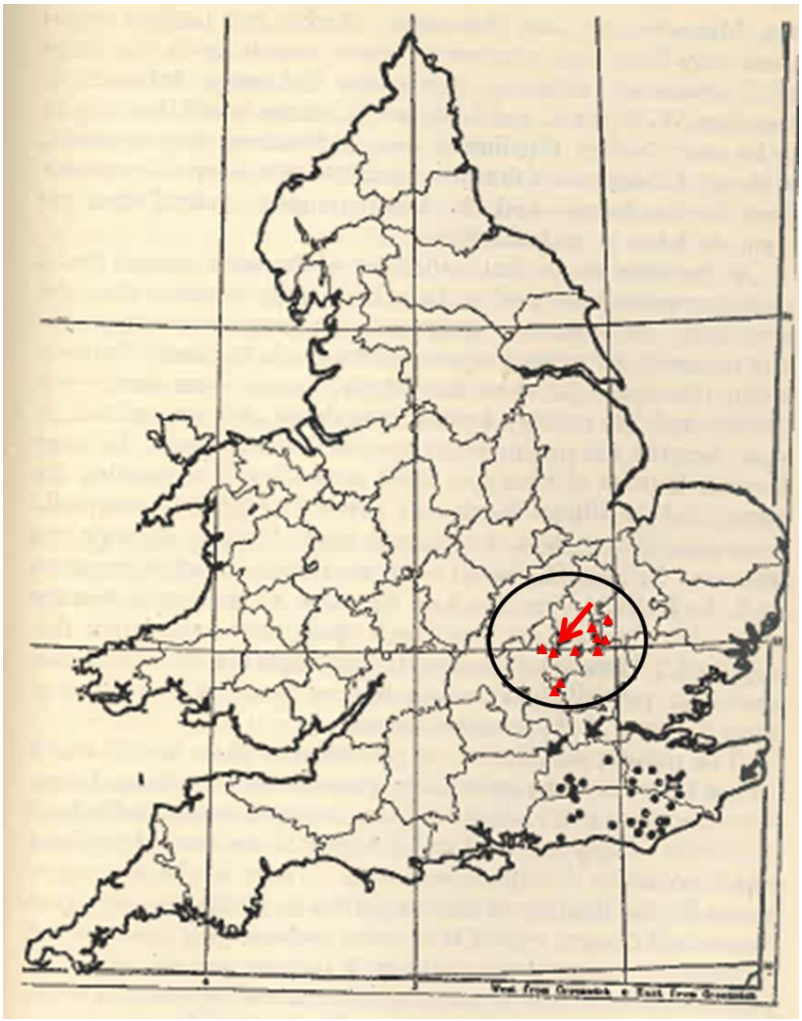


Figure 1.5. The distribution of melanic grey squirrels (*Sciurus carolinensis*) in England and Wales in 1954

Melanic squirrels are represented as red triangles within the black oval. The red arrow within the oval shows the probable introduction point. Dots and arrows outside the oval refer to albino squirrels. Map adapted from Shorten (1954).

The latest data on the spread of the black squirrel in Britain is currently being collected from the public via the website www.blacksquirrelproject.org in a project led by the author

(McRobie, 2014a). This website was launched in 2012 on BBC 1's The One Show by the author and it allows members of the public to record sightings of wildtype and melanic grey squirrels, as well as red squirrels, on an interactive map of the British Isles. The results are recorded visually, as shown in figure 1.6, taken from the website (McRobie, 2014a).

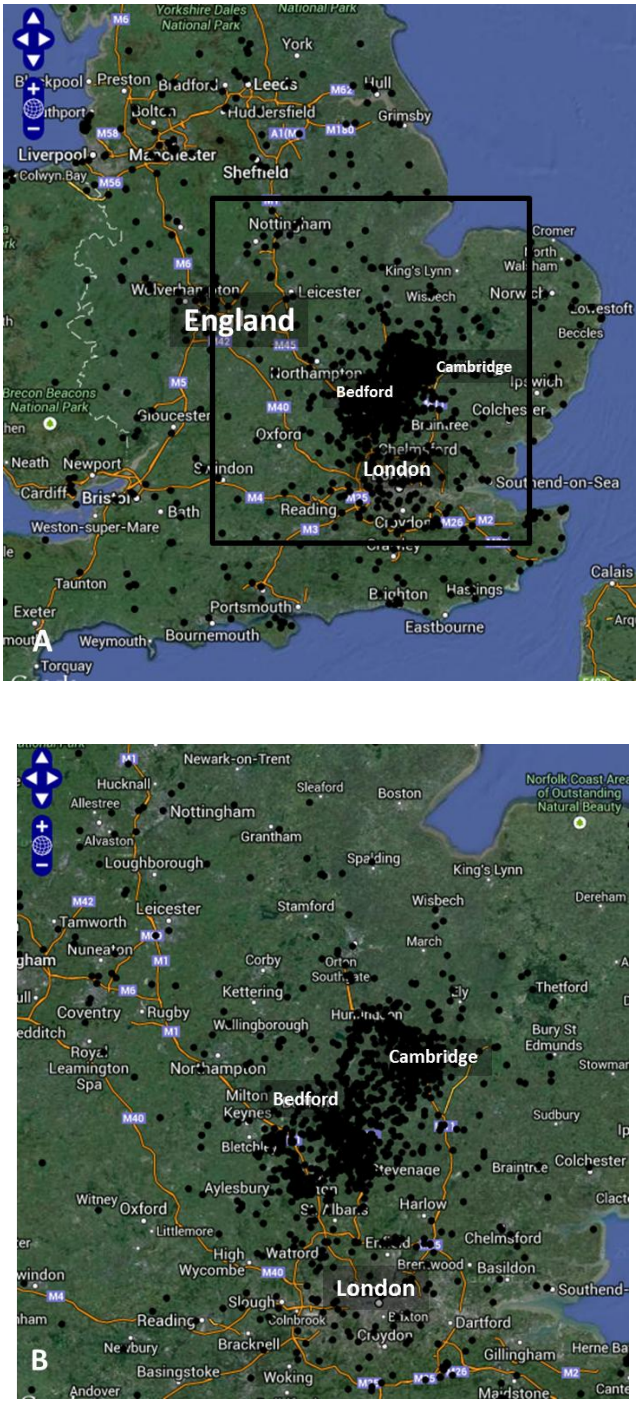


Figure 1.6. Map of melanic grey squirrel sightings in southern England in 2014

A) Overview of melanic grey squirrel sightings across southern England. The boxed area is shown as a close up in B) showing the most concentrated sightings around Bedfordshire, Hertfordshire and Cambridgeshire recorded on the www.blacksquirrelproject.org website. Maps adapted from McRobie (2014a).

This map shows that the majority of black squirrel sightings in the British Isles are in Bedfordshire, Hertfordshire and Cambridgeshire but also shows sightings throughout the south of England. Over 6000 sightings have been recorded to date (2014); however, there are clear limitations to this method of data collection. The website has been more widely advertised, via local radio, television and newspapers, in regions where black squirrels are known to live, which may lead to an ascertainment bias, however, the vast majority of sightings were recorded within hours of a national broadcast. Also the data is likely to be over-represented in densely populated areas and under-represented in sparsely populated areas. At present, there is no way of verifying the sightings and the data have not been collected in a strictly scientific way. Despite these shortcomings, the data does, however, give an interesting overview of the population of black squirrels in Britain.

1.3 Melanism

Many diverse species of mammals have both a wildtype and a melanic variant, for example rabbits (*Oryctolagus cuniculus*), grey seals (*Halichoerus grypus*), pigs (*Sus scrofa*), jaguars (*Panthera onca*) and mice (*Mus musculus*) (Robbins et al., 1993; Kijas et al., 1998; Majerus, 1998; Eizirik et al., 2003). The compound responsible for pigmentation in mammals is melanin which is ubiquitous in the animal kingdom, being found in such diverse organisms as yeast, fossilized dinosaurs and the ink of the squid (Bond, 2013). Melanin has a variety of functions including thermoregulation (Barrowclough and Sibley, 1980), protection (Goldstein et al., 2004), signalling (Bokony et al., 2003; Jawor and Breitwisch, 2003; Roulin, 2004; Mundy, 2005) and crypsis (Burtt Jr, 1986; Majerus, 1998). A growing number of studies have shown that many organisms have evolved cryptic colouration which minimises visual detection by hunting predators (Vignieri, Larson and Hoekstra, 2010). The most obvious example of this is where there is a correlation between pigmentation and substrate colour. A classic example of how natural selection has affected pigmentation is demonstrated in the case of the peppered moth (*Biston betularia*) (see figure 1.7) in the British Isles. Peppered moths are typically light coloured and speckled such that they match the lichen found on trees in their natural habitat. Industrialisation led to a change in this habitat as trees were no longer covered in lichen, but covered in soot and were darker as a consequence (Majerus, 1998). Frequencies of the melanic variant of this moth, now being camouflaged on the darker trees, increased due to selective predation by birds (Kettlewell, 1955; Majerus, Brunton and Stalker, 2000; Cook and Saccheri, 2013).

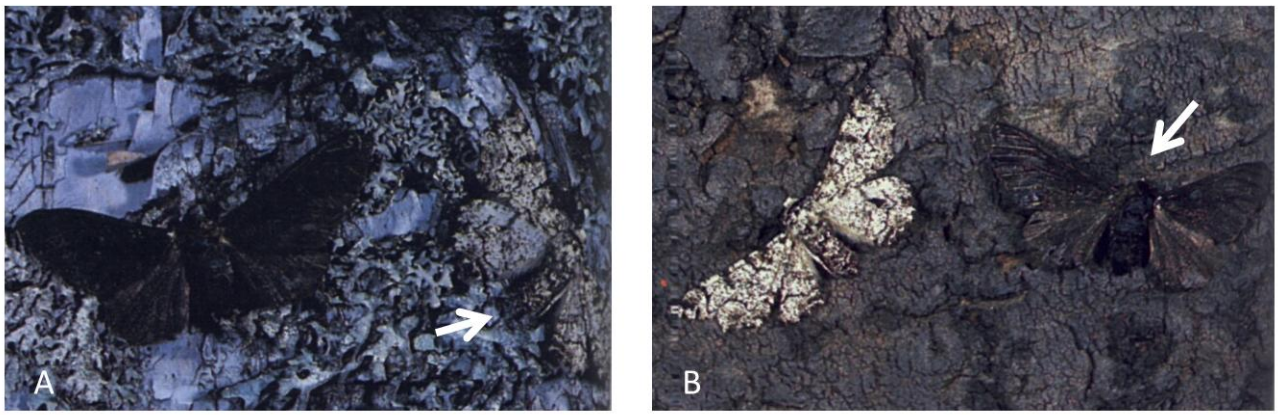


Figure 1.7. Wildtype and melanic morphs of the peppered moth (*Biston betularia*) showing cryptic colouration

A) Melanic and wildtype moths on light tree bark and B) on dark tree bark. White arrows indicate the more concealed morph in each case. Figure adapted from Boecklen (2014).

There are many other cases emerging where there is a correlation between pigmentation and substrate colour, for example in beach mice (*Peromyscus polionotus*). The mice that inhabit areas with dark, loamy soils have dark dorsal coats whereas mice that live on sandy dunes have a significantly lighter coat colour (Steiner, Weber and Hoekstra, 2007). A similar situation is found in pocket mice (*Chaetodipus intermedius*) and figure 1.8 shows the different morphs on different substrates demonstrating cryptic colouration (Nachman, Hoekstra and D'Agostino, 2003). Furthermore, it has been demonstrated that such cryptic colouration leads to less predation in mice, showing that such camouflage can indeed be adaptive (Vignieri, Larson and Hoekstra, 2010).

Figure 1.8. Different colour morphs of the pocket mice (*Chaetopidus intermedius*) on different coloured rocks demonstrating cryptic colouration

A) and D) show light mice on light and dark substrates and B) and C) show dark mice on light and dark substrates. Figure adapted from Nachman, Hoekstra and D'Agostino, (2003).

1.4 Mammalian Pigmentation

A great deal has been learnt about the genetic and molecular basis of colouration in a number of species of birds, fish, reptiles and mammals over the last hundred years (Lin and Fisher, 2007). The most comprehensive understanding has been gained through large scale genetic and mutagenesis studies of mice, which have revealed that over 100 loci and nearly 1000 alleles are involved in pigmentation (Miltenberger et al., 2002; Bennett and Lamoreux, 2003; Steingrímsson, Copeland and Jenkins, 2006).

1.4.1 Melanocytes

The pigment melanin is a heterogeneous, insoluble, non-protein biopolymer of no fixed molecular weight which is synthesised in specialised cells called melanocytes. Mammalian melanocytes derive from the neural crest in the early embryo. Melanocytes migrate to the epidermis, the iris and choroid of the eye and the inner ear. Follicular melanocytes are derived from epidermal melanocytes and reside in the hair bulb where they form part of the follicular-melanin unit. This unit consists of hair matrix melanocytes and keratinocytes (see figure 1.9). In humans there is typically one melanocyte for every five keratinocytes (Mayer, 1973; Bennett, 1993; Slominski and Paus, 1993; Goding, 2000; Tobin and Paus, 2001; Slominski et al., 2005b).

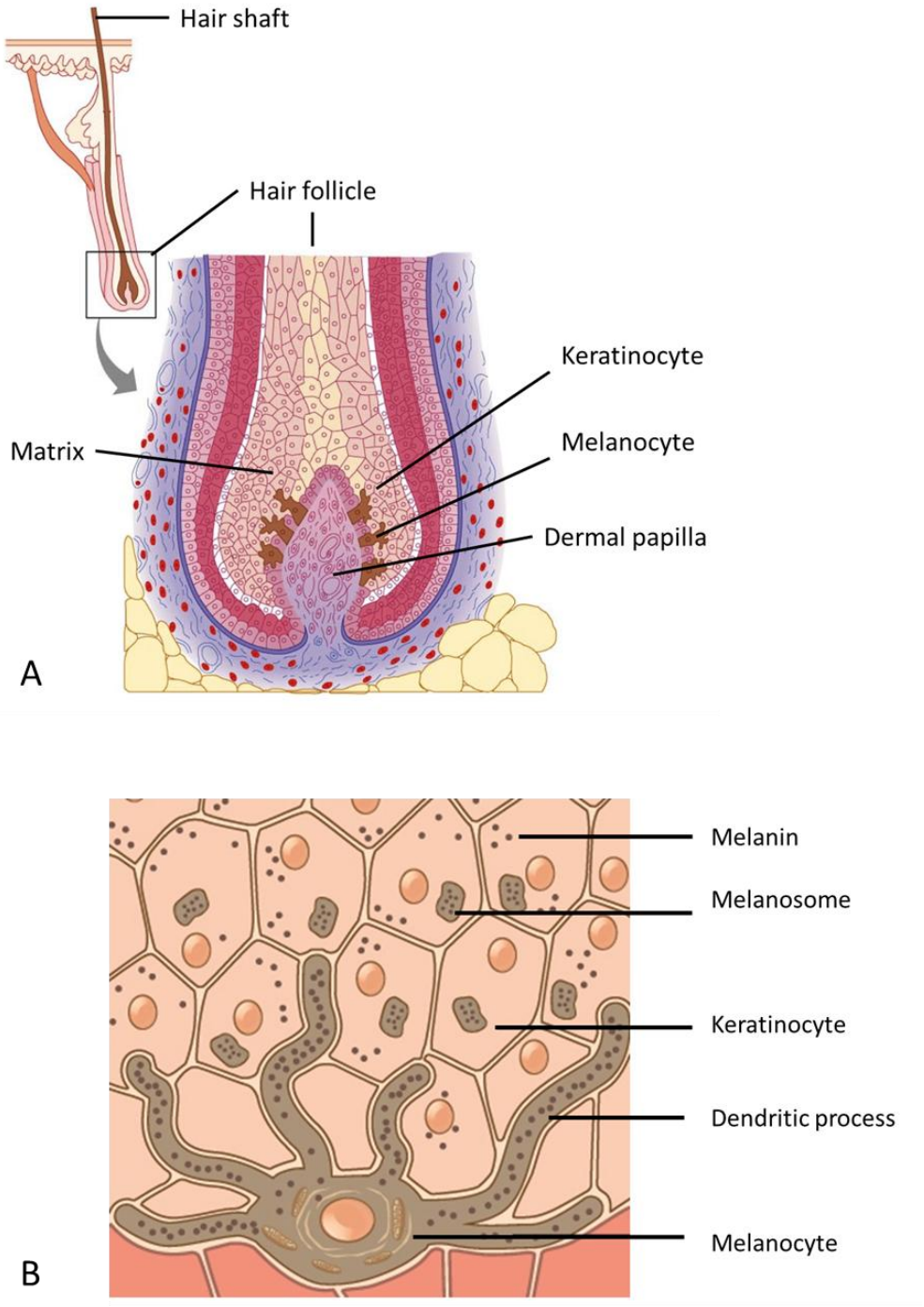


Figure 1.9. Melanocytes of the hair follicle

A) Hair follicle showing the position of melanocytes in the matrix. Figure adapted from Marieb (2007). B) Melanocyte with surrounding keratinocytes. Figure adapted from Hedegard (2014).

From now on, only melanocytes of the hair follicle will be considered, unless otherwise stated. In mice, the precursors of melanocytes are melanoblasts which arise from the neural crest at approximately embryonic day eight. Once in the hair follicle, the melanoblasts either differentiate into melanocytes or become melanocyte stem cells which form new melanocytes during each hair cycle (Wehrle-Haller and Weston, 1995; Nishimura et al., 2002; Steingrímsson, Copeland and Jenkins, 2006). Melanocytes have a number of dendritic processes which interdigitate between surrounding keratinocytes (see figure 1.9). There are many key genes involved in melanocyte development, including the paired box 3 gene (*PAX 3)* which is a transcription factor influencing proliferation, migration and differentiation (Kubic et al., 2008), and the microphthalmia transcription factor (*MITF*) which is a master regulator of melanocyte development, function and survival (Goding, 2000), Another key gene is the v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog or "*KIT"* gene which is involved in melanoblast expansion, survival and migration (Wehrle-Haller, 2003). Mutations in any of these genes can lead to white spotting where no melanocytes, and thus no pigments, are found (Lin and Fisher, 2007).

1.4.2 Melanin synthesis

Within melanocytes, melanin synthesis takes place in melanosomes, which are discrete membrane-bound organelles derived from late endosomes from the endoplastic reticulum. Melanosomes have four distinct stages of development. Stage one melanosomes have irregular arrays of amyloid fibrils which subsequently become the matrix (Slominski et al., 2004). Stage two melanosomes are segregated from the endocytic pathway and the fibrils are organised into arrays of parallel sheets which form the mature matrix. The enzymes

tyrosinase (TYR) and tyrosinase related protein-1 (TRP-1) are delivered to stage two melanosomes and melanin synthesis begins (Slominski et al., 2005b). Melanin is deposited on the fibres and by stage four the melanosome is saturated with melanin and internal structures are obscured (see figure 1.10).

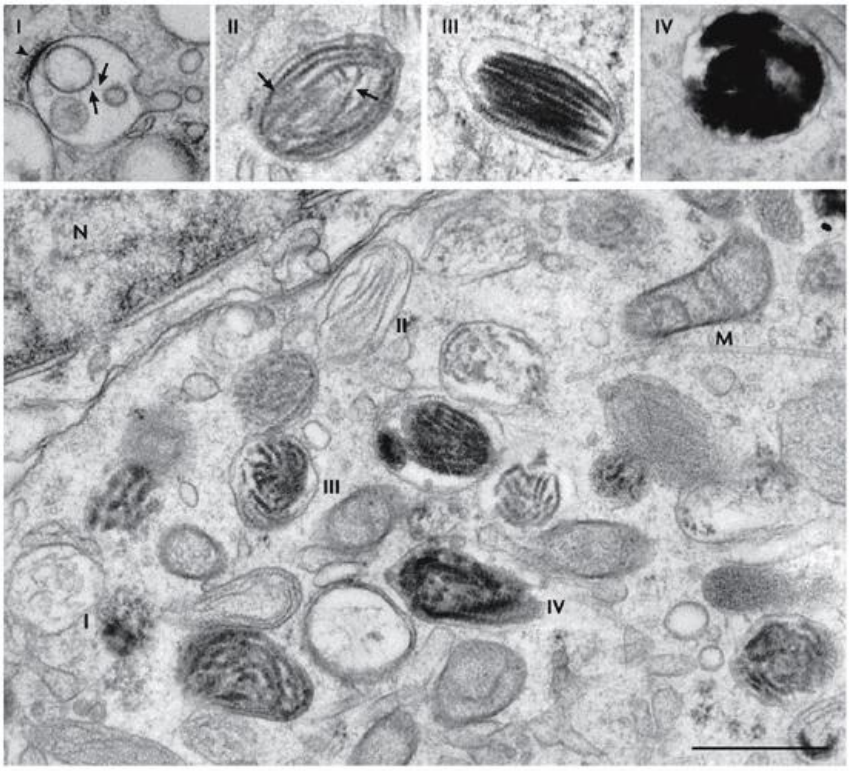


Figure 1.10. Electron microscopy of melanosomes

The four stages of melanosome development are highlighted. Scale bar represents 0.5 μm. M= mitochondria. Figure taken from Raposo and Marks (2007).

Saturated melanosomes are then transported along the dentritic processes and transferred to neighbouring keratinocytes of the hair matrix (see figure 1.11) (Ortonne and Prota, 1993). It has been demonstrated that filopodia from melanocyte dendrites act as conduits for

melanosome transfer and it has been suggested that dendricity is an important regulator of melanin transfer to keratinocytes (Scott et al., 2002; Slominski et al., 2005b). Melanin is ultimately transferred to the hair shaft in the keratinocytes. However, melanin synthesis only occurs during certain periods during the life cycle of a hair. The hair growth cycle is well characterised and the activity of a hair follicle is divided into three distinct phases: resting (telogen), growth (anagen) and regression (catagen) (Chase, 1954) and melanin synthesis only occurs during the anagen phase of this cycle (Slominski et al., 1994).

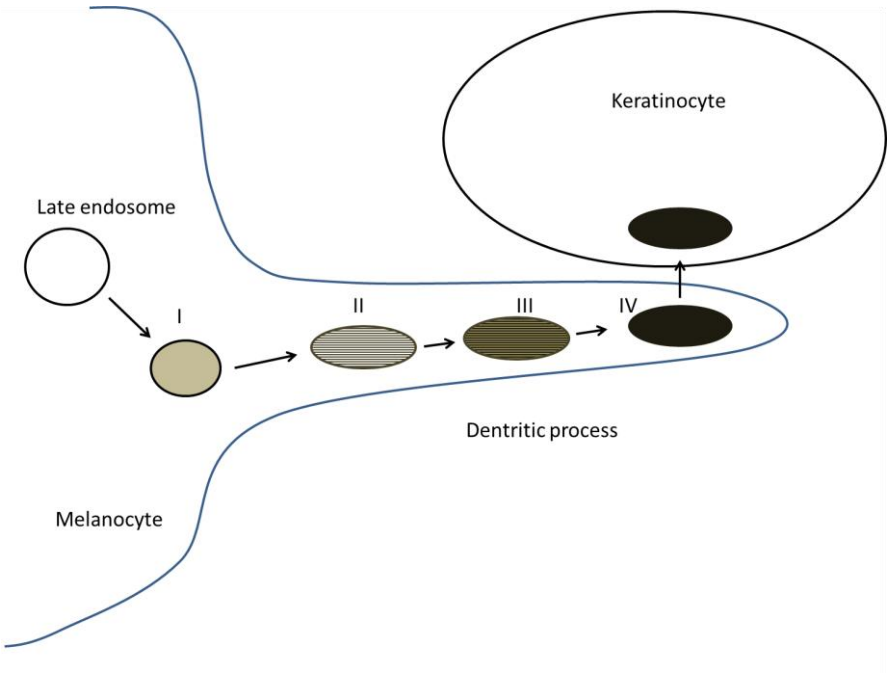


Figure 1.11. Melanin development and transfer to keratinocytes

Representation of a melanocyte, showing melanosomes containing melanin, in four stages of development, followed by transfer to a neighbouring keratinocyte along a dentritic process.

1.4.3 Eumelanin and phaeomelanin

Mammals produce two types of melanins; black/brown eumelanins and red/yellow phaeomelanins. Eumelanosomes are large and ellipsoidal with a highly ordered matrix, in contrast to phaeomelanosomes which are small and spherical with a disordered matrix (see figure 1.12). A single melanocyte can produce both eumelanosomes and phaeomelanosomes and they can co-exist in the same cell. These two distinct melanosomes do not, however, share a melanogenic pathway. Black hair melanosomes contain the largest number of dense eumelanosomes, whereas brown hair melanosomes are smaller and contain less melanin while blonde hair melanosomes contain very little melanin (Øyehaug et al., 2002). Although eumelanin and phaeomelanin have separate biochemical pathways, there is evidence that eu- and phaeomelanins can co-polymerise and produce variable shades of colour (Ozeki et al., 1995; Slominski et al., 2004). The rate-limiting step for synthesis of both types of melanin is the hydroxylation of tyrosine to L-3, 4-dihydroxyphenylalanine (L-DOPA) and oxidation of L-DOPA to DOPAquinine; these reactions are catalysed by TYR (Hearing and Tsukamoto, 1991; Hearing, 1999). Further enzymes are required for the synthesis of eumelanin: TRP-1 and dopachrome tautomerase (DCT). These enzymes are up-regulated in the presence of high levels of cyclic adenosine monophosphate (cAMP) suggesting that cAMP levels are pivotal in the switch from eumelanogenesis to phaeomelanogenesis (Busca and Ballotti, 2000; Øyehaug et al., 2002). Eumelanin is derived from DOPAchrome whereas phaeomelanins are derived from 5-S-cysteinylDOPA, as summarised in figure 1.12 (Pawelek, 1991; del Marmol and Beermann, 1996; Sturm, Teasdale and Box, 2001).

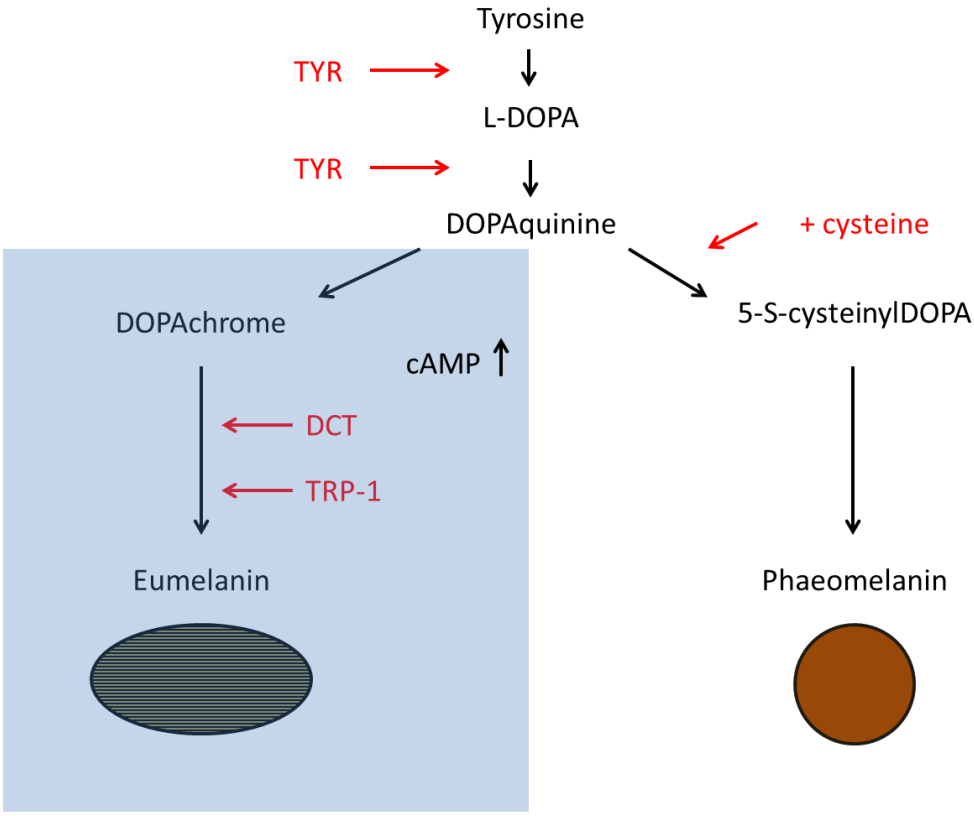


Figure 1.12. Summary of melanin biosynthesis pathways

Tyrosine is hydroxylated to L-DOPA and oxidation of L-DOPA produces DOPAquinine, catalysed by tyrosinase (TYR). In the presence of high levels of intracellular cAMP, DOPAchrome tautomerase (DCT) and tyrosine related protein (TRP-1) further modify DOPAquinine to produce eumelanin, shown in the blue area. Low levels of intracellular cAMP, and the addition of cysteine, lead to the production of phaeomelanin (*Pawelek, 1991*)..

1.4.4 Pigment type switching

Melanocytes are able to switch from producing eumelanin to producing phaeomelanin and it is this switch that is central to determining the final hair colour of an organism. The protein product of the *extension* locus has been identified as a principal regulator in the production of these two pigments. This locus encodes the MC1R which acts as a molecular switch and is

central to pigment type switching. When the agonist, alpha-melanocyte stimulating hormone (α-MSH) binds to the MC1R, the receptor becomes activated, which leads to raised intracellular levels of cAMP by coupling to adenylate cyclase activity (Cone et al., 1993; García-Borrón, Sánchez-Laorden and Jiménez-Cervantes, 2005). However, when the inverse agonist, agouti signalling protein (ASIP) binds to the MC1R, the receptor becomes inactivated and intracellular cAMP levels fall (Abdel-Malek et al., 2001). High levels of cAMP lead to eumelanogenesis and low levels lead to phaeomelanogenesis. When levels of cAMP become critically low, all pigment production ceases as summarised in figure 1.13 (Abdel-Malek et al., 2001).

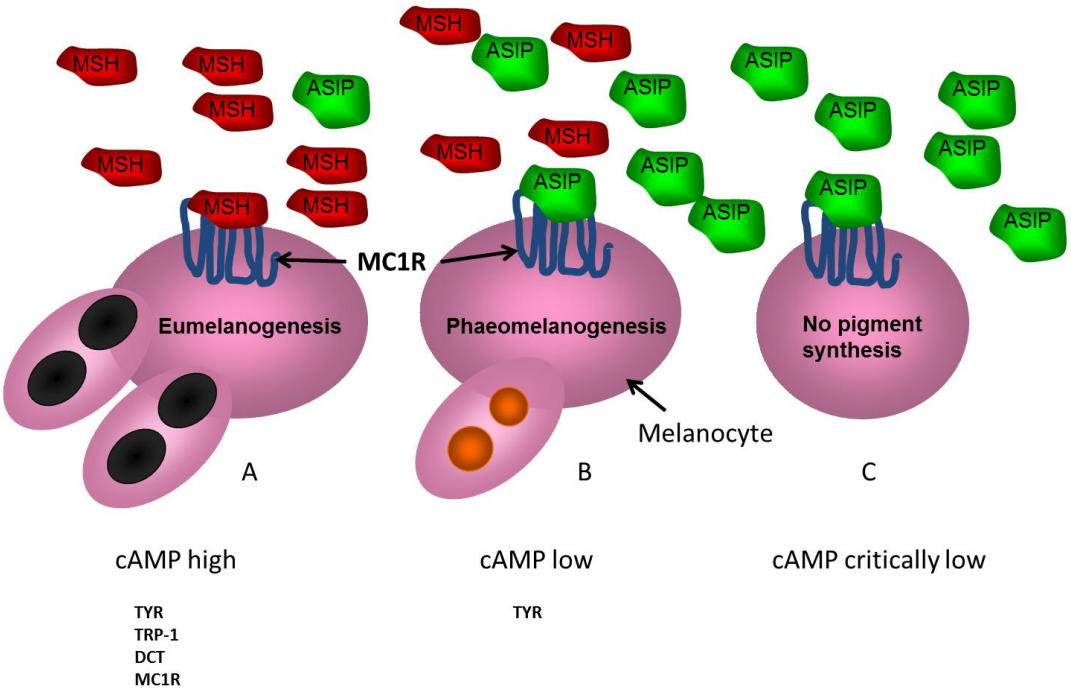


Figure 1.13. Schematic representation of pigment switching in melanocytes

A) The agonist, α-melanocyte stimulating hormone (α-MSH) binds to the melanocortin-1 receptor (MC1R) leading to raised cAMP levels in melanocytes and up-regulation of tyrosinase (TYR), tyrosinase related protein 1 (TRP-1), dopachrome tautomerase (DCT) and MC1R leading to eumelanogenesis. B) The inverse agonist, agouti signalling protein (ASIP) binds to the MC1R *leading to lowered cAMP levels in melanocytes and down-regulation of TRP-1, DCT and MC1R leading to phaeomelanogenesis. C) Continued treatment with ASIP leads to critically low levels of cAMP leading to down-regulation of TYR, TRP-1, DCT and MC1R resulting in cessation of all pigment production (Slominski et al., 2004).*

Thus, cAMP levels are a critical and pivotal aspect of melanocyte activity. The levels of TYR, TRP-1 and DCT determine which type of pigment is synthesised and these levels are ultimately determined by cAMP levels in the cell. cAMP binds to protein kinase A (PKA) allowing its catalytic subunit to become activated. Once activated, PKA phosphorylates its substrates which include ion channels, enzymes and regulatory proteins. In the nucleus, PKA

phosphorylates the cAMP response element binding protein (CREB) family of transcription factors and CREB-binding protein (CBP) (Slominski et al., 2004). Phosphorylated CREBs and CBPs are required for the expression of specific PKA-dependent genes. A key transcription factor that is activated by raised cAMP levels is the MITF of the *microphthalmia* locus (Hodgkinson et al., 1993; Hughes et al., 1993). MITF is involved in transcription of the *TRP-1, TYR* and *DCT* genes necessary for eumelanogenesis (Bertolotto et al., 1996; Bertolotto et al., 1998). Lowered levels of cAMP lead to lowered levels of these key enzymes and consequently a switch to phaeomelanogenesis. MITF is also involved in transcription of *MC1R* and so a positive feedback loop may be established where increased stimulation of MC1R leads to raised cAMP levels which in turn up-regulates the expression of MC1R (Aoki and Moro, 2002). The sequence of events described here is summarised in figure 1.14.

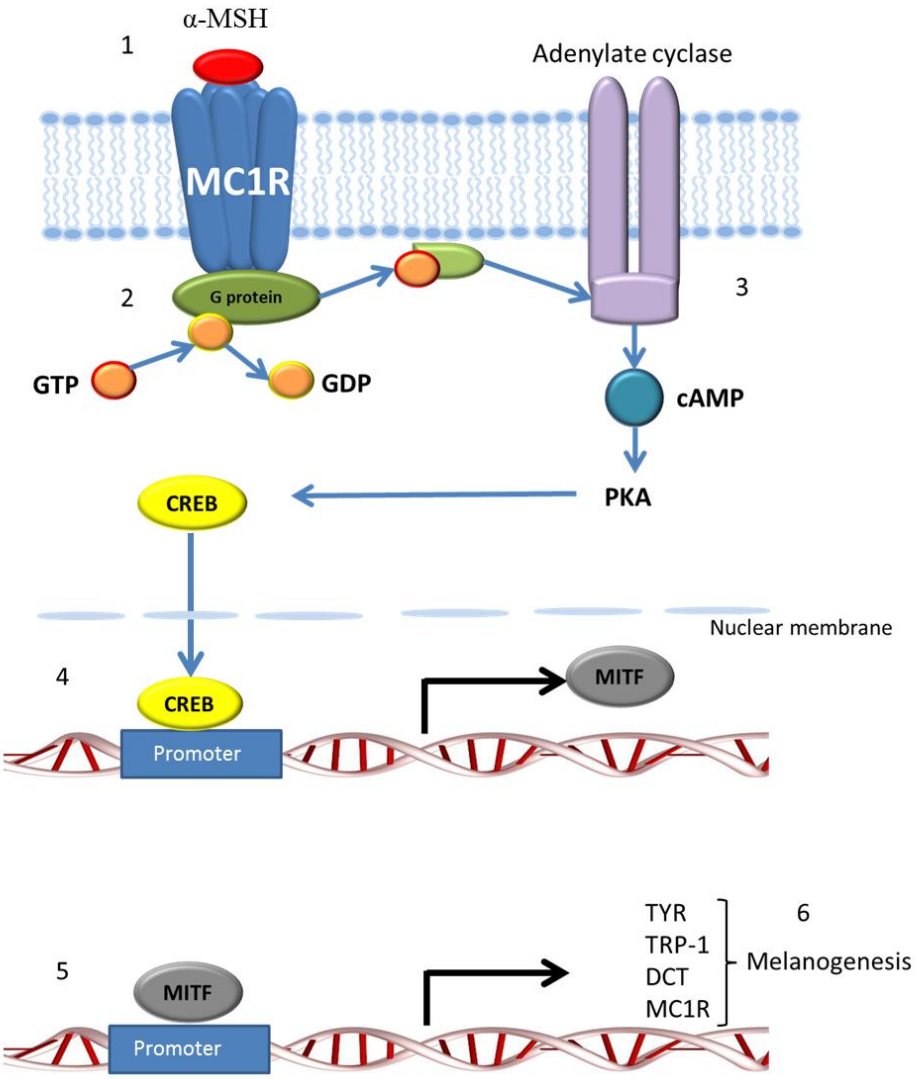


Figure 1.14. Schematic representation of the key regulatory pathways involved in melanogenesis

1) α-melanocyte stimulating hormone (α-MSH) binds to the melanocortin-1 receptor (MC1R) leading to receptor activation. 2) Activation of the MC1R induces the α-subunit of a G_S protein to release GDP, bind to GTP and dissociate into α- and βγ-subunits. The α-subunit bound to GTP stimulates adenylate cyclase. 3) Adenylate cyclase produces cAMP which binds to protein kinase A (PKA). PKA becomes activated and phosphorylates the cAMP response element binding protein (CREB) transcription factor. 4) CREB binds to the promoter region of the MITF gene on the DNA leading to the expression of MITF. 5) MITF binds to the promoter regions of tyrosinase (TYR), tyrosinase related protein-1 (TRP-1) dopachrome tautomerase (DCT) and MC1R genes. 6) TYR, TRP-1, DCT and MC1R are expressed leading to melanogenesis (Slominski et al., 2004).

As well as up-regulating genes involved in eumelanogenesis, cAMP is involved in increased dendricity of melanocytes by inhibiting the small GTP-binding protein, Rho, which then disrupts actin organisation (Busca et al., 1998; Busca and Ballotti, 2000). Rho inhibition also enhances cAMP stimulation of *TYR* gene transcription and thus eumelanogenesis (Ridley and Hall, 1992; Ridley and Hall, 1994). Raised cAMP levels are also associated with increased cell survival and proliferation (Pawelek, 1979; Pawelek, 1985; Rouzaud et al., 2003). Taken together, raised levels of cAMP have the overall effect of increased production of eumelanin.

1.5 The Genetics of Pigmentation

1.5.1 The *melanocortin-1 receptor* gene

The MC1R is a principal regulator of pigmentation, as described above. The *MC1R* gene was first cloned by Chhajlani and Wikberg (1992) and Mountjoy et al., (1992) and was first mapped to the *extension* locus by Robbins et al., (1993). The *MC1R* gene is intronless, with 945 coding base pairs (bp) giving a 314 amino acid protein in mice (Mountjoy et al., 1992) but this number can vary from species to species. The gene is predicted by Rouzaud et al. (2003) to have a regulatory region of 2400 bp upstream of the gene, as shown in figure 1.15.

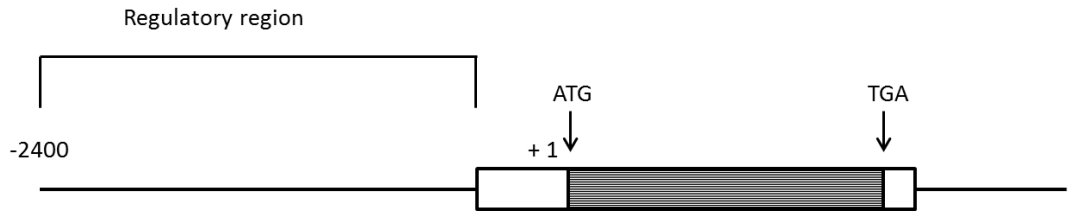


Figure 1.15. Genomic organisation of the melanocortin-1 receptor (*MC1R*) gene

The horizontal line indicates DNA (not to scale). The exon of the MC1R is shown as a box, with horizontal stripes indicating the coding region (945 bp). The untranslated regions of the exon are shown as unfilled boxes. Vertical arrows indicate the start and stop codons. The numbers refer to base pairs where +1 refers to the A of the start codon and -2400 to the upstream base pairs of the regulatory region (Rouzaud et al., 2003).

This regulatory region of the *MC1R* is predicted to have many potential binding sites for transcription factors, including 30 consensus regulatory elements for specificity protein-1 (SP1), five for activator protein-1 (AP1) and 10 for activator protein-2 (AP2) as well as 13 enhancer boxes (E boxes) (Rouzaud et al., 2003). The gene encodes a guanine nucleotide

binding protein-coupled receptor (GPCR) predominantly expressed in melanocytes (Donatien et al., 1992) and a general schematic representation of the MC1R is shown in figure 1.16.

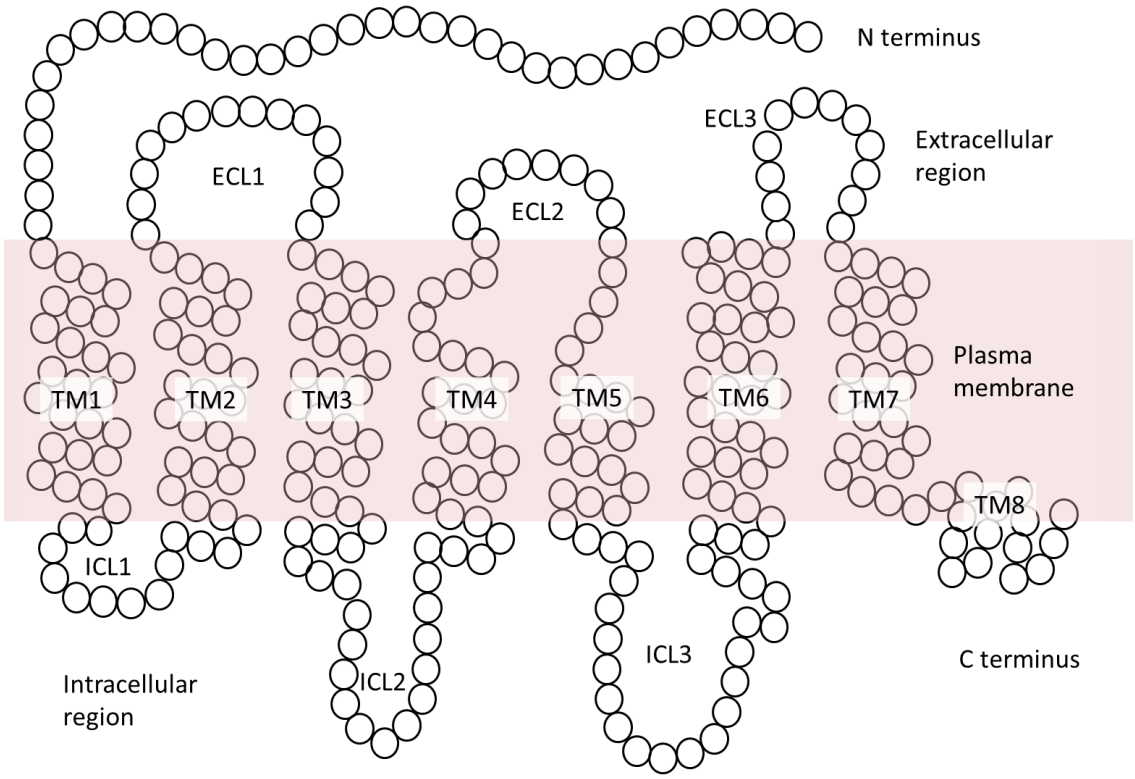


Figure 1.16. Schematic representation of the melanocortin-1 receptor

Circles represent amino acids and the pink rectangle represents the plasma membrane. TM= transmembrane helix, ICL= intracellular loop and ECL= extracellular loop.

1.5.2 The melanocortin-1 receptor and pigmentation

The central role of the MC1R as a key regulator of pigmentation has been revealed by extensive studies on pigmentation and mutants in mice (Slominski et al., 2004). Wildtype mice have dark hairs with a subapical band of phaeomelanin, but there are many variations

from this. The MC1R exhibits a high level of constitutive activity in mice. A receptor is considered to be constitutively active if it can be activated in the absence of an agonist. This constitutive activity alone is sufficient to support eumelanogenesis, as mice lacking the pro-opiomelanocortin (*POMC*) gene from which the agonist α-MSH is derived, are indistinguishable from normal mice (Slominski et al., 2005a). Interestingly this is not the case in humans as a lack of *POMC* leads to red hair. This is because MC1R is expressed ten times more in mice than humans (Jackson et al., 2007). MC1R-deficient mice produce solid yellow hairs and ASIP-deficient mice produce solid black hairs. Therefore, a functional MC1R and ASIP both appear to be necessary for full pigment switching to take place in mice, and presumably this is also the case for other animals.

1.5.3 Mutations in the *melanocortin-1 receptor* gene associated with melanism

Melanism has been associated with mutations in the *MC1R* gene in a number of mammals and birds (Majerus and Mundy, 2003; Mundy, 2005). Such mutations are usually gain-of-function mutations and are inherited in a dominant or incompletely dominant way. The majority of mutations involve an amino acid substitution and these are generally located in the transmembrane helices. A hotspot for mutations is found at the extracellular side of the second and third transmembrane helices (TM) with 11 out of 24 substitutions and 4 out of 5 deletions clustering in this region, as shown in figure 1.17 and 1.18.

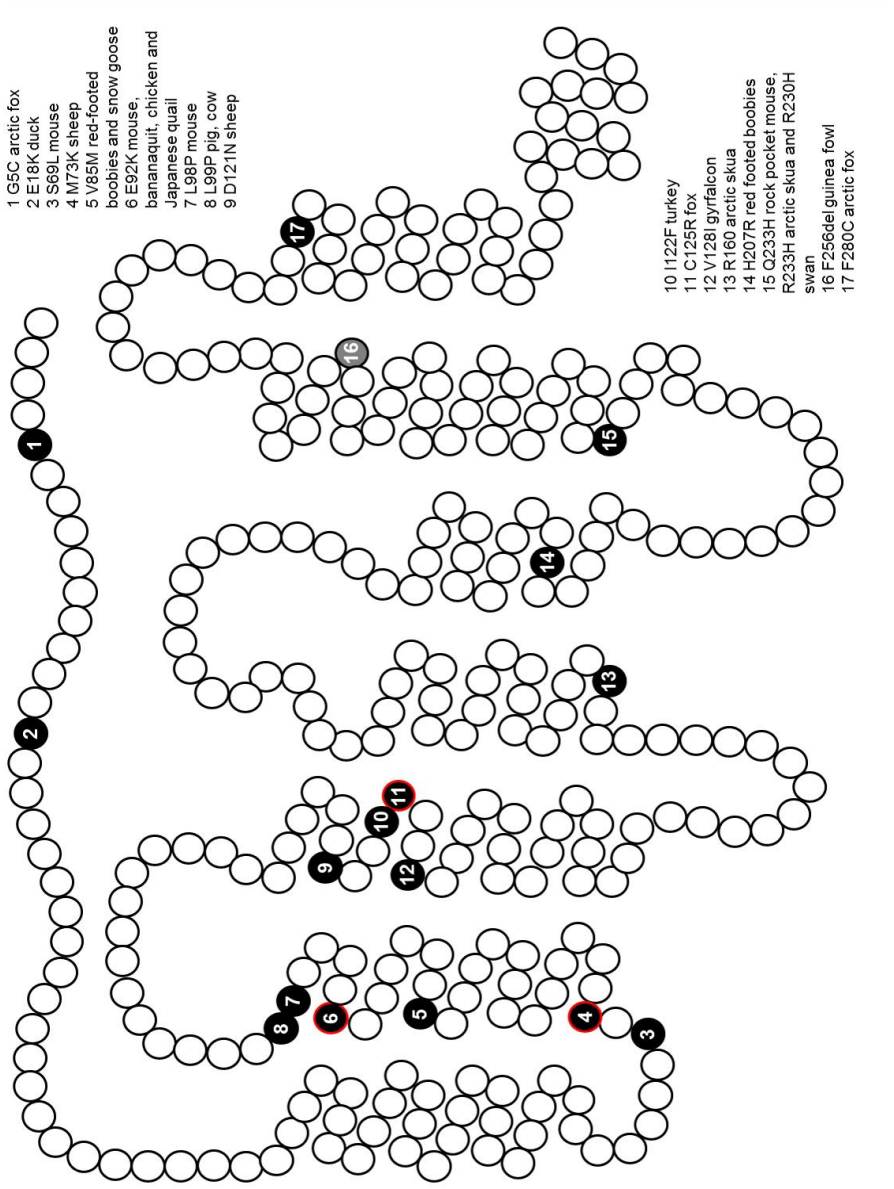


Figure 1.17. Schematic representation of the melanocortin-1 receptor showing mutations associated with melanism in different species

Circles represent amino acids, black circles with white numbering represent substitutions associated with melanism, the grey circle with white lettering is a deleted amino acid associated with melanism and circles with red outlines represent substitutions that are known to lead to a constitutively active receptor (Robbins et al., 1993; Våge et al., 1997a; Våge et al., 1999; Theron et al., 2001; Nachman, Hoekstra and D'Agostino, 2003; Mundy, 2005; Våge et al., 2005; Nadeau, Minvielle and Mundy, 2006; Baião, Schreiber and Parker, 2007; Pointer and Mundy, 2008; Vidal et al., 2010; Vidal, Viñas and Pla, 2010; Johnson, Ambers and Burnham, 2012; Yu et al., 2013).

The melanic "*sombre*" mouse has a Lys at position 92 replacing the negatively charged Glu (E92K) (mouse numbering) and this is located at the extracellular side of TM2 (Robbins et al., 1993). This dominant mutation, associated with melanism, is also found independently in three other species: the bananaquit (*Coereba flaveola*), the chicken (*Gallus gallus*) and the Japanese quail (*Coturnix japonica*) (Theron et al., 2001; Kerje et al., 2003; Nadeau, Minvielle and Mundy, 2006). Interestingly, this mutation is also found in the ruffed lemur (*Varecia variegate*) but does not appear to cause complete melanism (Mundy and Kelly, 2003; Pointer and Mundy, 2008). A different mutation, changing a Leu to a Pro at position 98 (L98P) in mice is associated with melanism. This is located at the extracellular side of TM2 (Robbins et al., 1993). The similar L99P mutation is found in cattle (*Bos taurus*) and pigs (Klungland et al., 1995; Kijas et al., 1998). A mutation changing an Asp to an Asn at position 121 (D121N) is associated with melanism in pigs and sheep (*Ovis aries*). This is positioned at the extracellular side of TM3 (Kijas et al., 1998; Våge et al., 1999). The D121N mutation in sheep is accompanied by a change from Met to Lys at position 73 (M73K) (Våge et al., 1999). The change from Cys to Arg at position 125 (C125R) is associated with melanism in the fox (*Vulpes vulpes*) and is also at the extracellular side of TM3 (Våge et al., 1997a). Other amino acid substitutions associated with melanism are R230H in swans (*Cygnus*) and arctic skuas (*Stercorarius parasiticus*) (Mundy et al., 2004; Pointer and Mundy, 2008), Q233H in pocket mice (Nachman, Hoekstra and D'Agostino, 2003), V128I in the gyrfalcon (*Falco rusticolus*) (Johnson, Ambers and Burnham, 2012), E18K in ducks (*Anas platyrhynchos*) (Yu et al., 2013), I122F in turkeys (*Meleagris gallopavo*) (Vidal, Viñas and Pla, 2010), V85M in the lesser snow goose (*Anser c. caerulescens*) (Majerus and Mundy, 2003; Mundy et al., 2004), H207R in red-footed boobies (*Sula sula*) (Baião, Schreiber and Parker, 2007) and G5C and F280C in the arctic fox (*Alopex lagopus*) (Våge et al., 2005) and E91A together with

E100K in black-necked swans (*Cygnus melanocoryphus*). Further mutations with less dramatic darkening are found in reindeer (*Rangifer tarandus*), M73T together with F280C (Våge et al., 2014) and the great skua (*Catharacta skua*), R230C (Mundy et al., 2004).

1.5.4 Mutations in the *melanocortin-1 receptor* gene which lead to constitutively active receptors

A number of the mutations in the MC1R, described above, have been investigated further and found to lead to a constitutively active receptor, meaning that the receptor has a higher level of activity in the absence of agonist binding. Mutagenesis and functional studies on the E92 position in mice, chickens and lemurs reveal that a change from an acidic to basic amino acid, either lysine or arginine, leads to constitutive activity (Robbins et al., 1993; Ling et al., 2003; Haitina et al., 2007). Binding studies show, however, that the E92K reduces α-MSH binding (Robbins et al., 1993; Haitina et al., 2007). The C125R mutation in foxes and the M73K mutation in sheep also lead to constitutive activation (Våge et al., 1997a; Våge et al., 1999). All of these mutations introduce a positive charge to the receptor and it is hypothesised that this charge mimics the effect that agonist binding has on the receptor (Lu, Vage and Cone, 1998).

1.5.5 Deletions in the *melanocortin-1 receptor* gene associated with melanism

A number of mutations associated with melanism involve deletions and these deletions are generally found at the extracellular side of TM2, as shown in figure 1.18.

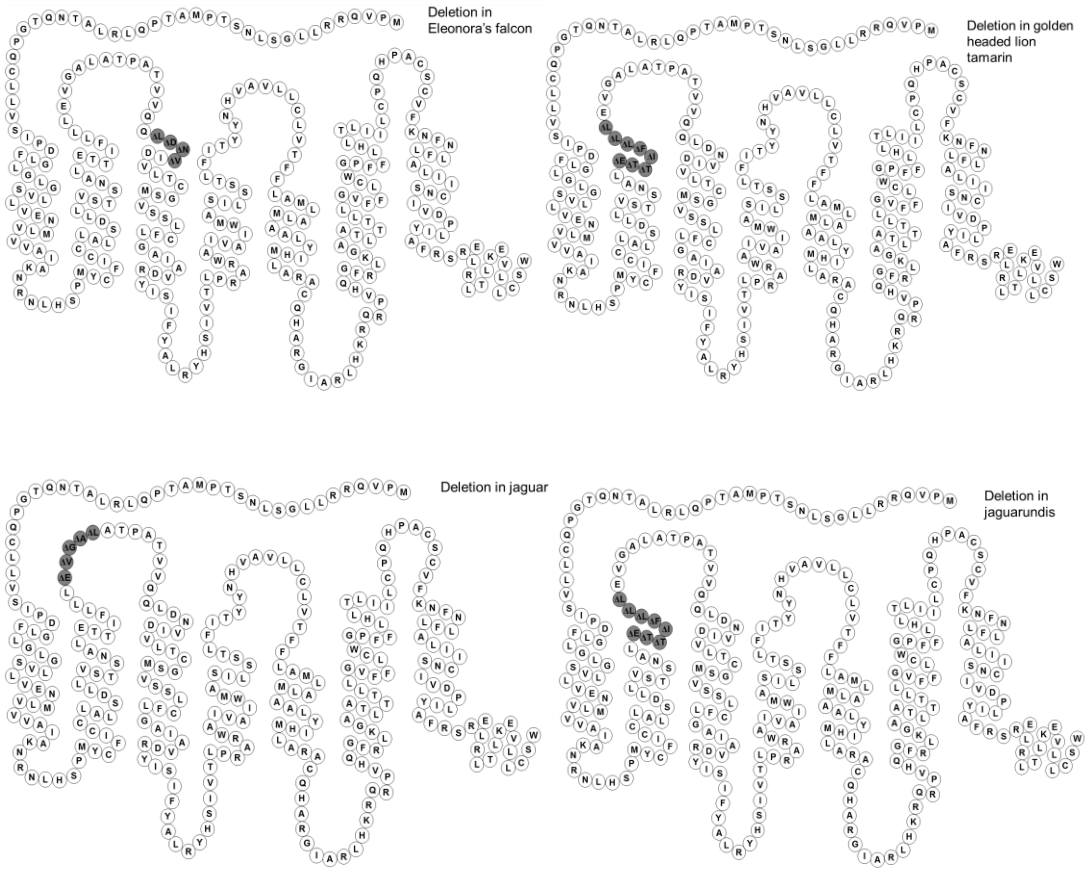


Figure 1.18. Schematic representation of the melanocortin-1 receptor showing deletions associated with melanism

*Circles represent amino acids and dark grey circles represent amino acids deleted from transmembrane 2 and extracellular loop 1 region associated with melanism in Eleonora's falcon (*Falco eleonorae*), the golden-headed lion tamarin (*Leontopithecus chrysomelas*), jaguar (*Panthera onca*) and jaguarundi (*Herpailurus yaguarondi*) (Eizirik et al., 2003; Mundy and Kelly, 2003; Gangoso et al., 2011).*

Melanism in jaguars is associated with a 15 bp deletion and in the jaguarundi (*Herpailurus yaguarondi*) with a 24 bp deletion (Eizirik et al., 2003). A similar 24 bp deletion has been found in the golden-headed lion tamarin (*Leontopithecus chrysomelas*), although the face of the tamarin is not melanic and causation has not been established (Mundy and Kelly, 2003).

A 12 bp deletion is associated with melanism in Eleonora's falcon (*Falco eleonorae*) in a similar position (Gangoso et al., 2011) and a 6 bp deletion is associated with melanism in rabbits (*Oryctolagus cuniculus*) (Fontanesi et al., 2006). A 3 bp deletion in a different position, near the extracellular side of TM6 is associated with melanism in the guinea fowl (*Numida meleagris*) (Vidal et al., 2010). Not all deletions at the extracellular side of TM2 lead to melanism. A number of such deletions have no apparent effect; for example a 15 bp deletion in the wolverine (*Gulo gulo*) , a 28 bp deletion in the stone marten (*Martes foina*), a 45 bp deletion in four species of martens (*Martes americana, M. martes, M. melampus and M. zibellina*) (Hosoda et al., 2005). A further 21 bp deletion has been reported in the gopher (*Thomomys bottae*) but this is in the N-terminal domain of the protein and is not associated with melanism (Wlasiuk and Nachman, 2007). However, a 30 bp deletion leads to a phaeomelanic phenotype in rabbits (Fontanesi et al., 2006). Many other mutations lead to phaeomelanic phenotypes.

1.5.6 Mutations in the *melanocortin-1 receptor* associated with lighter phenotypes

Mutations leading to phaeomelanic phenotypes include many red hair alleles in humans, for example the V60L, D84E, V92M, R151C, I155T, R160W, R163Q and D294H mutations (Beaumont et al., 2005) and a S83F mutation leading to the chestnut coat colour in horses (*Equus caballus*) (Marklund et al., 1996). A Y298C mutation leads to a white coat colour in the Kermode bear (*Ursus americanus*) (Ritland, Newton and Marshall, 2001), a R112H mutation leads to phaeomelanic plumage in the brown booby (*Sula leucogaster*) (Baiao and Parker, 2012), a premature stop codon leads to yellow coat colour in dogs (*Canis lupus familiaris*) (Everts, Rothuizen and Oost, 2000) and a R65C mutation leads to a lighter

phenotype in beach mice. This mutation in beach mice leads to a decrease in receptor signalling and consequent reduction in eumelanogenesis (Steiner, Weber and Hoekstra, 2007). Three independent mutations in the *MC1R* of three species of white lizards, the eastern fence lizard (*Sceloporus undulates*) (H208Y), little striped whiptail (*Aspidoscelis inornata*) (T170I), and lesser earless lizard (*Holbrookia maculate*) (V168I) are statistically associated with lighter phenotypes (Rosenblum et al., 2010).

Melanism is not always associated with mutations in the *MC1R*. Melanism in wall lizards (*Podarcis lilfordi* and *P. pityusensis*), Japanese ornamental carp (*Cyprinus carpio*) and the blue-crowned manakin (*Lepidothrix coronate*) is not associated with amino acid changes in the MC1R. However, this does not completely rule out the involvement of this gene as regulatory regions were not investigated in these cases (Cheviron, Hackett and Brumfield, 2006; Bar et al., 2013; Buades et al., 2013).

1.5.7 Mutations in *agouti signalling protein* associated with melanism

ASIP, as described above, is an inverse agonist to the MC1R and a number of natural and induced mutants have revealed much about its functions. Typically, loss-of-function mutations of ASIP are recessive and lead to melanism, whereas gain-of-function mutations are dominant and lead to lighter phenotypes. Many mutations leading to melanism are found in the three coding exons of *ASIP*. In the Asian golden cat (*Pardofelis temminckii*), a non-synonymous mutation changing a Cys to a Trp (C124W) in exon 4 is associated with recessive inheritance of melanism (Schneider et al., 2012). Recessive black phenotypes in German Shepherd dogs are associated with a missense mutation in exon 4 changing an Arg to

a Cys at position 96 (R96C) found in the C-terminal domain of the protein (Kerns et al., 2004). This extra cysteine seems to be more disruptive than loss of cysteines reported by extensive studies on laboratory mice (Miltenberger et al., 2002). An almost purely black phenotype in the black lion tamarin (*Leontopithecus chrysopygus*) is associated with a missense mutation in exon 4 changing a Lys to a Pro at position 89 (L89P) (Mundy and Kelly, 2006). A missense mutation in exon 4 of the Sulawesi macaque (*Macaca nigra*), changing a Pro to a Ser at position 102 (P102S) is associated with melanism (Mundy and Kelly, 2006). It has been suggested that the L89P and P102S mutations are not causative as a proline is found in position 89 in other species (mouse, cow, cat and fox) where no melanism is observed, and serine is found in position 102 in the fox (Våge et al., 1997a). However, the overall ASIP sequences of these other species differ considerably to the tamarin and macaque and it seems possible that these mutations could cause melanism. In leopards (*Panthera pardus*) a single nucleotide polymorphism causes a non-synonymous mutation in exon 4 which is predicted to introduce a premature stop codon (Schneider et al., 2012).

Many mutations in the *ASIP* gene leading to melanism are the results of deletions and insertions causing frameshifts. Melanism in foxes is associated with a 166 bp deletion which removes the start codon and entire signal peptide from *ASIP* (Våge et al., 1997b). This deletion is likely to result in the complete absence of functional ASIP and the dark phenotype is the same as that observed in foxes with a constitutively active MC1R. An 11 bp deletion in exon 2 of *ASIP* is associated with the recessive black coat colour of horses (Rieder et al., 2001). This deletion is out of frame and therefore would result in a non-functioning protein. Similarly, a 5 bp deletion has been identified in Soay sheep (*Ovis aries*), which produces a

premature stop codon at position 64 (Gratten et al., 2010). A 19 bp deletion in exon 2 is associated with melanism in rats (*Rattus norvegicus*). The deletion causes a premature stop codon where the shortened protein lacks the N-glycosylation site and the C-terminal cysteine residues essential to function (Kuramoto et al., 2001). Melanism in deer mice (*Peromyscus maniculatus*) is caused by two independent mutations in the *ASIP* gene. Here, a 125 kb deletion, which includes exons 1 and 2, results in a complete lack of *ASIP* expression. In a different population of the mice, a non-synonymous mutation changing glutamine to a stop codon in exon 3 leads to the elimination of exon 4 (Kingsley et al., 2009). In rabbits, an insertion of a single bp in exon 2 just downstream of the start codon creates a premature stop codon and a completely non-functional protein resulting in recessively inherited melanism (Fontanesi, et al., 2008; Fontanesi et al., 2010). In domestic cats (*Felis catus*), a 2 bp deletion in exon 2 causes a frameshift mutation and complete loss of the C-terminal domain. This is completely associated with melanism and is recessively inherited (Eizirik et al., 2003).

A number of natural mutations in regulatory regions of *ASIP* have also been identified which are associated with melanism. It was found that a silenced promoter for *ASIP* was associated with recessive melanism in sheep (Norris and Whan, 2008). A short interspersed element and a poly(A) stretch was found to be inserted within the *ASIP* promoter in the Boxer dog and is associated with deregulated *ASIP* expression and the darker black-and-tan phenotype (Ciampolini et al., 2013). It has been postulated that a regulatory region of the *ASIP* gene is responsible for the black-and-tan pigmentation in Mangalitza pigs (Drögemüller et al., 2006) and changes in pigmentation in cattle (Girardot et al., 2005).

1.5.8 Mutations in *agouti signalling protein* associated with lighter phenotypes

The majority of mutations in *ASIP* reported in the literature are loss-of-function mutations, as there are many more ways to break a gene than to make it function more effectively. However, there are a few cases of gain-of-function mutations leading to a lighter phenotype, for example, the missense mutations changing an Ala at position 82 to a Ser and an Arg at position 83 to a His (A82S and R83H) are associated with fawn or sable coats in dogs, which is a lighter phenotype to the black-and-tan and tricolor phenotype (Berryere et al., 2005). Mutations in *ASIP* associated with lighter phenotypes are generally found to be associated with different expression levels. Extensive laboratory studies have been carried out involving analysis of the phenotypes of mice that have undergone germ cell mutagenesis (Miltenberger et al., 2002). The results revealed that mutations in the *ASIP* gene with the most dramatic effects on phenotype are found in the C-terminal domain and that generally complete abrogation of agouti function was due to mutations in the coding regions. In contrast, mutations with milder effects are found elsewhere, in the regulatory regions of the gene, which alter expression levels (Miltenberger et al., 2002).

The relationship between phenotype and *ASIP* expression has been extensively researched in beach mice (*Peromyscus polionotus*) and deer mice. In beach mice lighter phenotypes are associated with an increase in *ASIP* mRNA (Steiner, Weber and Hoekstra, 2007) and in the deer mice levels of *ASIP* mRNA were higher and also the transcript was expressed for a longer period during the hair growth cycle resulting an overall lighter phenotype (Linnen et al., 2009).

1.5.9 Mutation in other genes associated with melanism

Other genes have been identified which have been associated with melanism, for example dominant black phenotypes of dogs are associated with a deletion of Gly at position 23 in β-defensin gene found at the *K* locus. This is a newly identified, high affinity ligand for the MC1R and behaves in a similar way to ASIP but does not function as an agonist or inverse agonist. It is suggested that the protein competitively inhibits ASIP and that the mutation could either have an increased affinity for MC1R or there may be an increased availability of the mature protein (Candille et al., 2007).

The MC1R forms the central focus of this study into the genetic and molecular basis of melanism in the grey squirrel. In order to understand the molecular mechanisms by which the MC1R might affect pigmentation, the general characteristics of GPCRs are considered next, followed by a consideration of the specific features of the MC1R structure and function.

1.6 G-protein Coupled Receptors

1.6.1 General features of G-protein coupled receptors

GPCRs, also known as seven transmembrane receptors, are proteins which span the cell's membrane and act as molecular switches at the interface between the cell's external and internal environment. GPCRs form one of the largest and most diverse membrane protein families with more than 800 genes, representing approximately 2% of the human genome (Fredriksson et al., 2003; Vassilatis et al., 2003). GPCRs have seven transmembrane domains with an extracellular amino terminus and intracellular carboxyl terminus. The generalised structure of a GPCR is shown in figure 1.19.

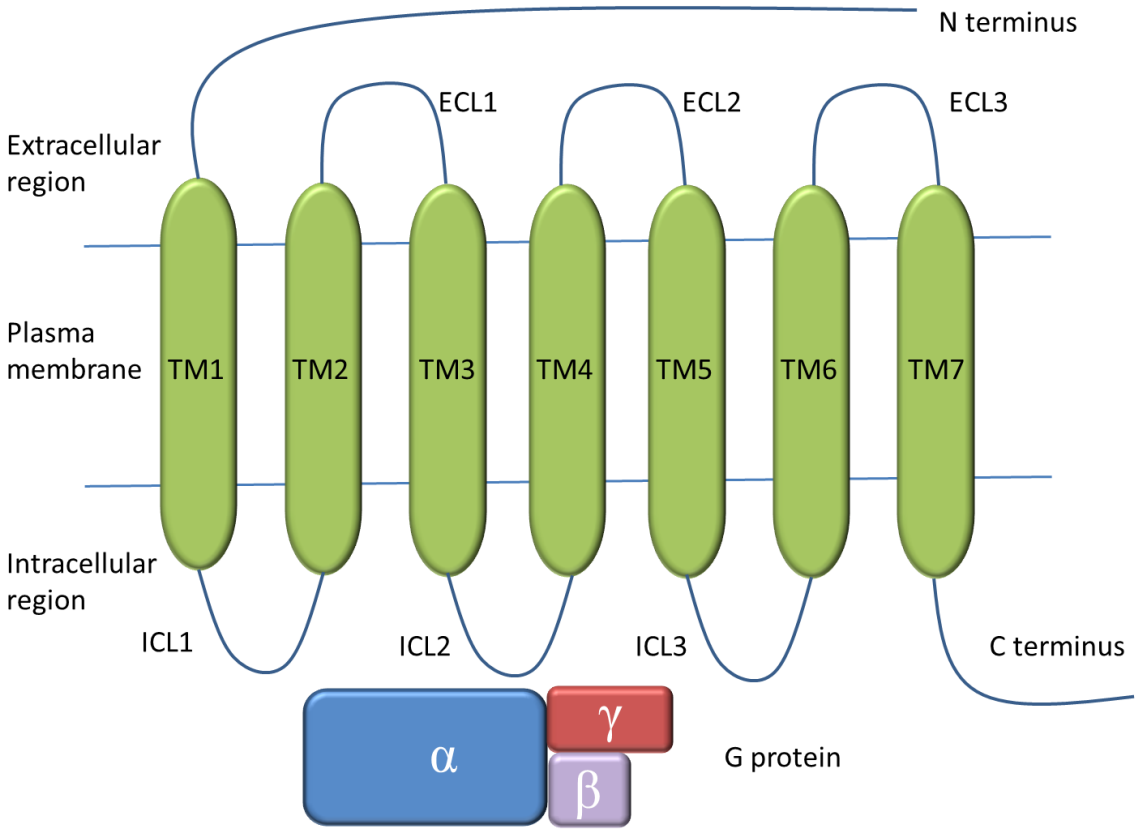


Figure 1.19. Schematic representation of a G-protein coupled receptor

A generalised G-protein coupled receptor (GPCR) showing seven transmembrane (TM) domains, three extracellular loops (ECL), three intracellular loops (ICL) and a G-protein. The G-protein has three subunits, α, β, and γ.

A diversity of ligands on the exterior of the cell membrane bind to GPCRs, including peptides, glycoproteins, lipids, ions and nucleotides and in the case of rhodopsin, photons are absorbed by the cofactor retinal in the transmembrane region. When a ligand binds to a GPCR, the receptor undergoes conformational changes which enable it to transduce signals across the membrane. GPCRs typically interact with intracellular heterotrimeric guanine nucleotide-binding (G) proteins (Kobilka, 2007). G-proteins are composed of three subunits; alpha, beta and gamma. An activated receptor induces the G-protein α subunit to release

guanosine diphosphate (GDP) and bind to guanosine triphosphate (GTP) and then dissociate from the βγ dimer (see figure 1.20) (Hamm, 1998; Kobilka, 2007). These subunits are then able to stimulate further intracellular cascades through downstream effectors. G-proteins can be divided into four broad families based on similarity of the alpha subunit. These include $G\alpha_{i/v'}$, $G\alpha_{s,}$ $G\alpha_{q/11',}$ $G\alpha_{12/13}$, and each family stimulates a distinct set of effectors (Cabrera-Vera et al., 2003). The G-protein primarily considered in this study is $G\alpha_{s,}$ which stimulates the effector adenylate cyclase, which goes on to catalyse the conversion of adenosine triphosphate (ATP) to cAMP (Knall and Johnson, 1998; Taussig and Zimmermann, 1998). GPCRs are also able to signal through other effectors, for example β arrestin (Lefkowitz and Shenoy, 2005; Galandrin, Oligny-Longpré and Bouvier, 2007). The exact mechanisms by which GPCRs work are not yet fully understood; however, a great deal of research has been carried out recently and much progress had been made. Indeed, the Nobel Prize for Chemistry in 2012 was awarded to Kobilka and Lefkowitz for "for studies of G-protein-coupled receptors".

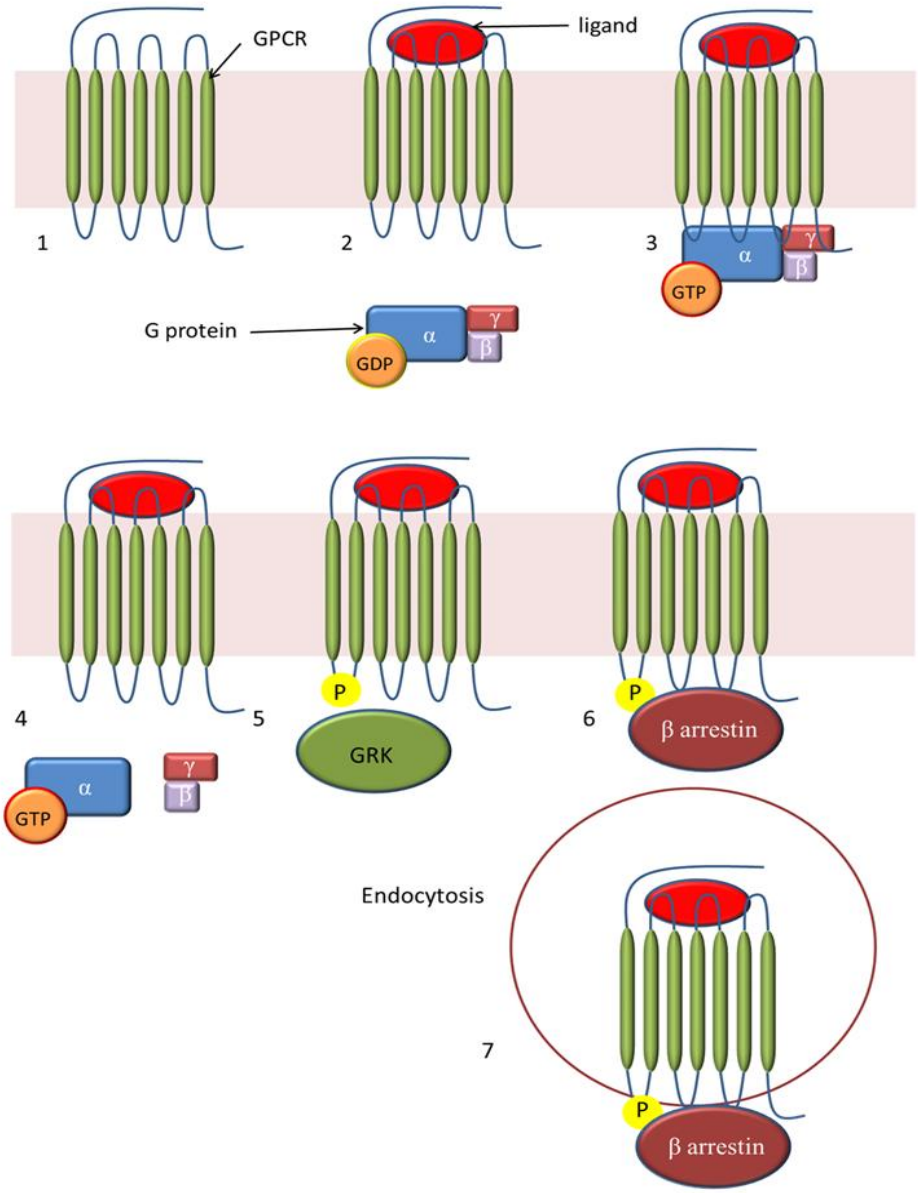


Figure 1.20. Summary of G-protein coupled receptor activation and desensitisation

1) G-protein coupled receptor (GPCR) with no ligand bound. 2) GPCR bound to an agonist on the extracellular side of the receptor. G-protein is bound to GDP. 3) When an agonist binds the G-protein binds on the intracellular side of the receptor and GDP is exchanged for GTP. 4) The G-protein subunits separate. 5) GRK phosphorylates the GPCR. 6) β-arrestin binds to the GPCR. 7) the GPCR is removed from the plasma membrane by endocytosis. GRK = GPCR kinase (Hamm, 1998; Cabrera-Vera et al., 2003).

1.6.2 Predicting the tertiary structure of G-protein coupled receptors

GPCRs are difficult to crystallise as they are intrinsically dynamic. The first high resolution structure of a GPCR was solved in 2000 with dark adapted bovine rhodopsin, where high resolution is generally considered to be below 3 Å (Beuming and Sherman, 2012). This 2.8 Å structure was elucidated using X-ray crystallography which allowed an accurate understanding of amino acid side chain conformations, extra- and intracellular loops and N- and C- terminal domains (Palczewski et al., 2000). Since 2000 many more GPCR structures have been solved and the results have revealed that overall folding of GPCRs varies very little. Since 2007, there have been many technical breakthroughs and innovations in crystallography. These techniques include creating receptor-T4 lysozyme and receptor-apocytochrome chimeras, co-crystallization with antibodies, thermostabilisation, engineering disulphide bridges and truncation of flexible regions (Venkatakrishnan et al., 2013). There are a number of drawbacks to these techniques; post-translational modifications are often lost and truncations and additions may affect the overall conformation of the receptor and give misleading structural information (Venkatakrishnan et al., 2013). Structures may be obtained in inactive conformations bound to an inverse agonist or antagonist or active conformations bound with agonists. These bound structures give a static snapshot of information on the structural changes taking place during receptor activation. The many crystal structures of GPCRs now available reveal that the overall folding is highly conserved and well predicted by rhodopsin (Venkatakrishnan et al., 2013). Figure 1.21 shows a typical three dimensional representation of a GPCR.

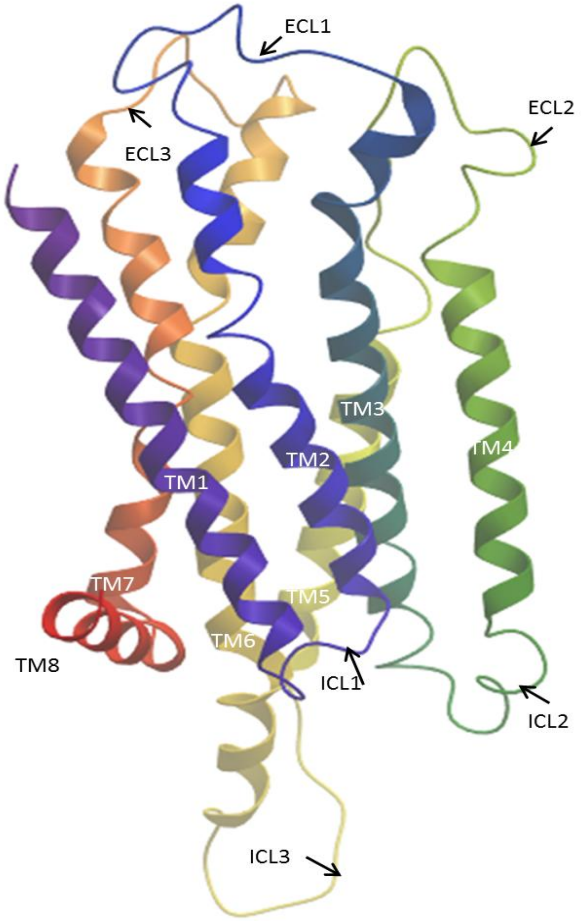


Figure 1.21. Three dimensional representation of a G-protein coupled receptor

Ribbon diagram representation of a typical G-protein coupled receptor with seven transmembrane helices, three extracellular loops, three intracellular loops and an amphipathic helix eight. The model is coloured using the rainbow spectrum from blue at the N- terminus to red at the C-terminus. TM = transmembrane helix, ICL = intracellular loop, ECL = extracellular loop.

1.6.3 Active and inactive states of G-protein coupled receptors

In simple terms, GPCRs can be thought of as bimodal switches with inactive and active states. The active state is the conformation which couples and stabilises an effector molecule, typically a G-protein (Gether, 2000). Most GPCRs have a basal level of activity in the absence of ligands due to their inherent conformational flexibility (Gether, 2000). This basal

activity means that the GPCR can activate the G-protein in the absence of an agonist. During basal activity, the GPCR is in equilibrium between R (inactive) and R* (active) states and the receptor is thought to alternate spontaneously from R to R*, even though this is both kinetically and conformationally complex. The level of basal activity is determined by the equilibrium between these two states and this can be influenced by pH, salt gradients, lipid bilayer composition and membrane potential (Deupi and Standfuss, 2011; Audet and Bouvier, 2012). Ligands can affect this basal activity and equilibrium in various ways. Agonists are ligands which bind to the receptor and stabilise the active R* conformation leading to full activation of a signalling pathway. Partial agonists lead to partial activation and antagonists have no effect on basal activity but competitively block other ligands. Inverse agonists bind to the receptor and stabilise the inactive R conformation which leads to a decrease in basal activity (see figure 1.22).

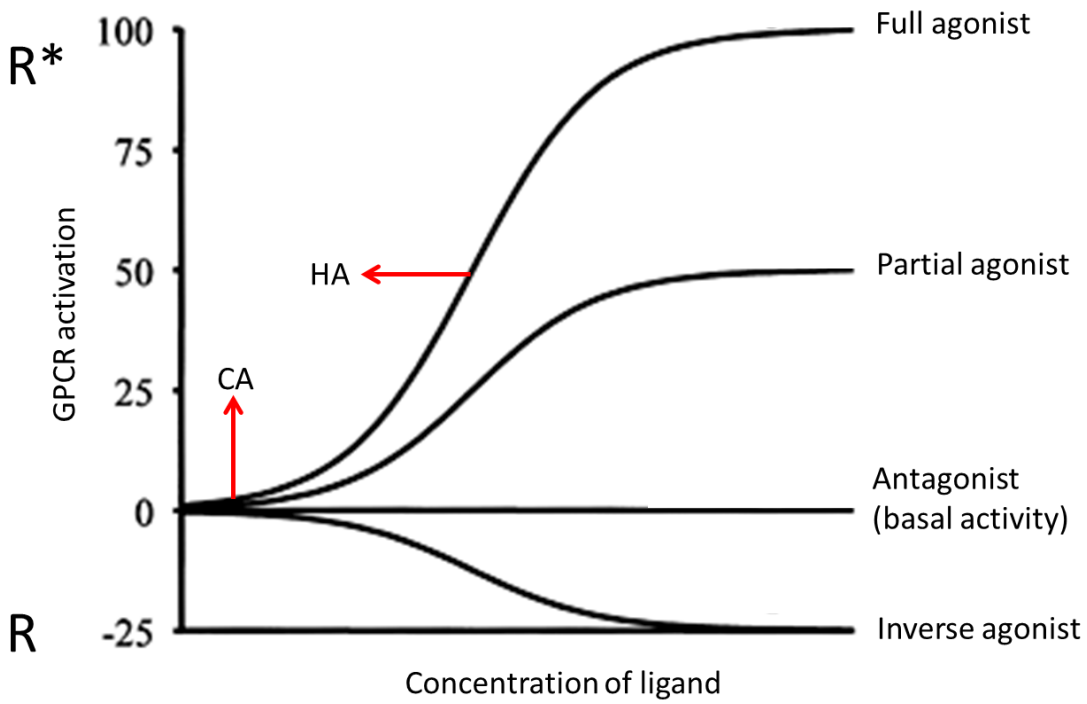


Figure 1.22. Summary of the effects of agonists, antagonists and inverse agonists on G-protein coupled receptor activity

Graph to illustrate the effects of different concentrations of ligands on G-protein coupled receptor activity. Full agonists lead to full activation (active R state), partial agonists lead to partial activation, antagonists have no effect on activity and inverse agonists lower the level of activation (inactive R state). The vertical red arrow indicates that the basal level of activity is increased for constitutively active (CA) receptors. The horizontal red arrow indicates that the graph is moved to the left for hyperactive receptors (HA), which require lower concentrations of ligand in order to be activated.*

1.6.4 Maintaining overall stability in G-protein coupled receptors

GPCRs are inherently flexible but in order to function effectively as switches it is important that the receptor has an overall stability. Analyses of diverse GPCRs have revealed a

consensus network of 24 inter-transmembrane contacts both in the active and inactive states (see figure 1.23) (Madabushi et al., 2004; Kobilka, 2007; Venkatakrishnan et al., 2013). Here, atoms are considered to be in contact if they are within the van der Waal interaction distance (Venkatakrishnan et al., 2013).

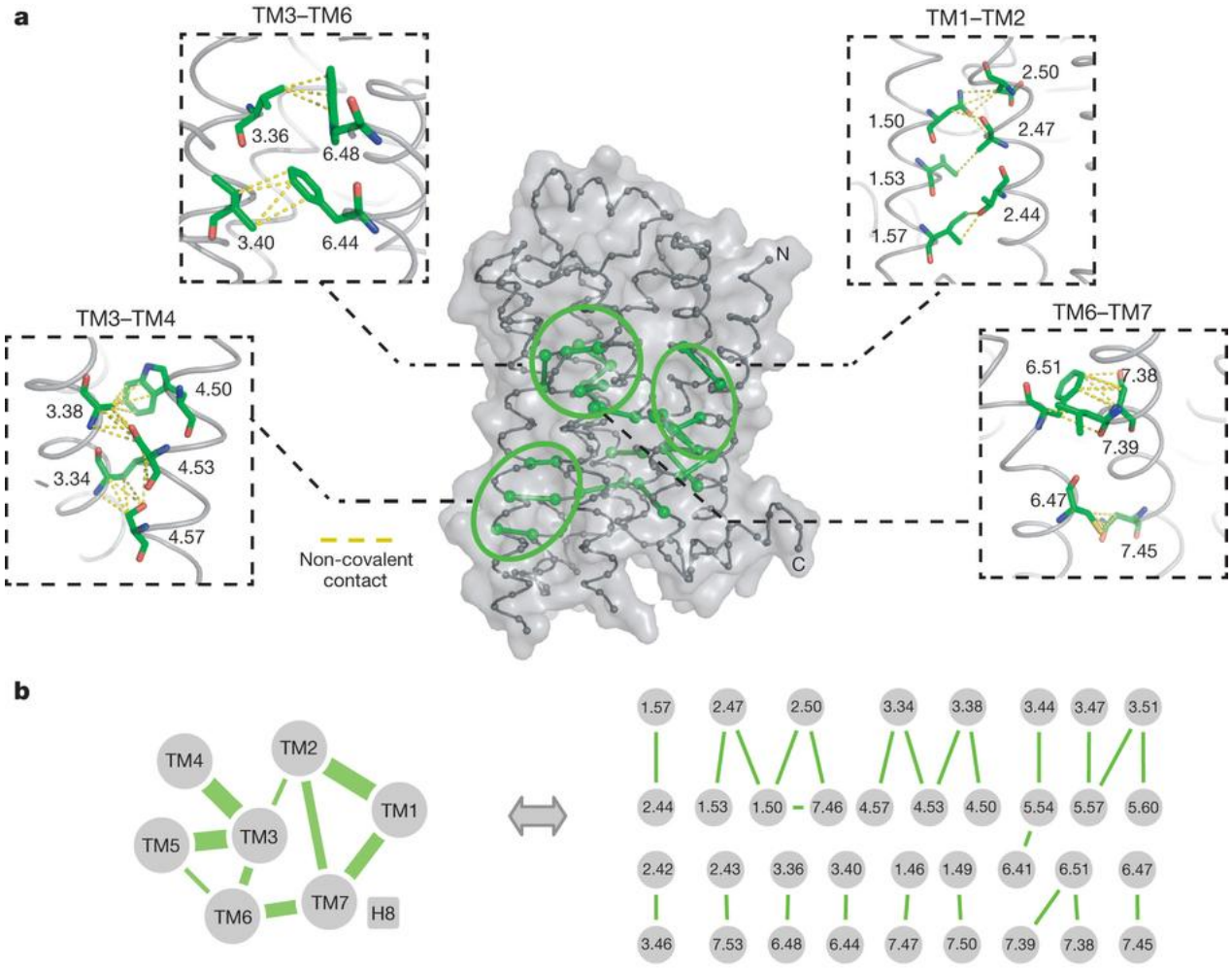


Figure 1.23. Consensus contacts between helices of G-protein coupled receptors

Network of 24 conserved inter-transmembrane contacts shown here on the β-adrenergic receptor. For clarity, contacts between transmembrane helices (TM) three and five are omitted. A) The β-adrenergic receptor with contacts highlighted with green ovals. Boxes show details of the contacts. Numbering is according to the Ballesteros-Weinstein system where the first number denotes the TM and the second number indicates the position relative to the most conserved residue of the helix which is given the number 50. Numbers lower than 50 are towards the N-terminus and numbers higher than 50 are towards the C-terminus. B) Schematic representation of the inter-TM contacts with TMs represented as circles. Lines represent the presence of consensus contacts. The thickness of the line is proportional to the number of contacts. Figure taken from Venkatakrishnan et al. (2013).

Given that the contacts are present in both R and R* states, it seems likely that the amino acids involved are structurally important residues, indeed, mutations in many of these positions affect receptor function, either increasing or decreasing it (Madabushi et al., 2004). Mutations which disrupt these non-covalent intramolecular interactions tend to increase the flexibility of the receptor and so increase the probability that the receptor will be in an active conformation (Madabushi et al., 2004). Transmembrane helix three (TM3) acts as an important structural and functional hub, making contact with four other helices. Almost every position of TM3 is vital for defining the fold of the receptor. At the cytoplasmic side of TM3 is the highly conserved Asp-Arg-Tyr "DRY" motif (shown later in figure 1.29). The Arg of this motif forms an important ionic lock with TM6 which stabilises the inactive conformation in some closely studied GPCRs. Indeed, when ionic locks are disrupted by reducing pH, a higher basal activity is observed (Ghanouni et al., 2001). One particular residue that is highly conserved seems to be vital in constraining intramolecular interactions: a change in a highly conserved Ala residue at the cytoplasmic side of TM6 to any other residue results in higher basal activity (Kjelsberg et al., 1992). This work highlights how constraining the molecule, and maintaining overall stability, are necessary for it to function effectively as a switch. Many other constraining intramolecular interactions have been found to be highly conserved, keeping the receptor silent in the absence of agonists (Gether, 2000).

1.6.5 Conformational changes in G-protein coupled receptors during activation

Structural studies of GPCRs in the active conformation have revealed that agonist binding induces a 2 Å shift of TM3 towards the extracellular side of the receptor (Venkatakrishnan et al., 2013). This movement provokes rearrangements of TM5 and TM6. Movement of TM6 involves a rotation and tilting relative to TM3 and this movement becomes amplified due to

the proline “kink” in TM6. The overall effect of this movement is to open the G-protein binding region on the cytoplasmic side of the receptor. It has been shown that TM6 moves by 14 Å in the β_2 adrenergic receptor following activation (Venkatakrishnan et al., 2013). Other studies have demonstrated that disruption of the ionic lock between TM3 and TM6 leads to the movement of TM6 away from the TM bundle creating a space for the G-protein and this space allows the DRY motif of TM3 to interact with the G-protein (Yao et al., 2006) (see figure 1.24).

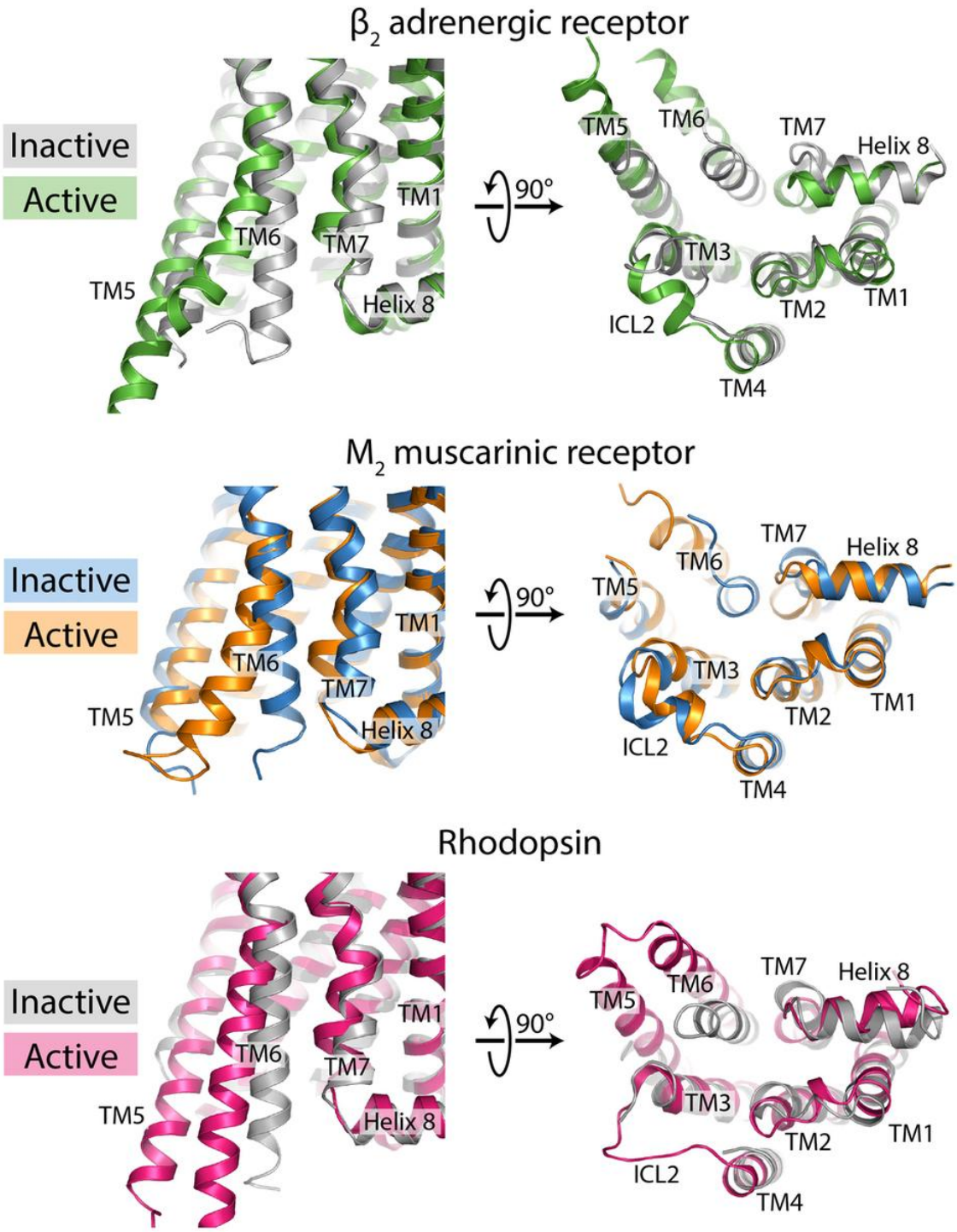


Figure 1.24. Active and inactive conformation of G-protein coupled receptors

The images to the left show a side view of the receptors, and the images to the right show the intracellular view of the receptors. Figure taken from Kruse et al. (2013).

Figure 1.25 shows a summary of the movement of TM6, illustrated in figure 1.24, that takes place during activation, illustrating the opening up of the G-protein binding cavity on the intracellular side of the receptor.

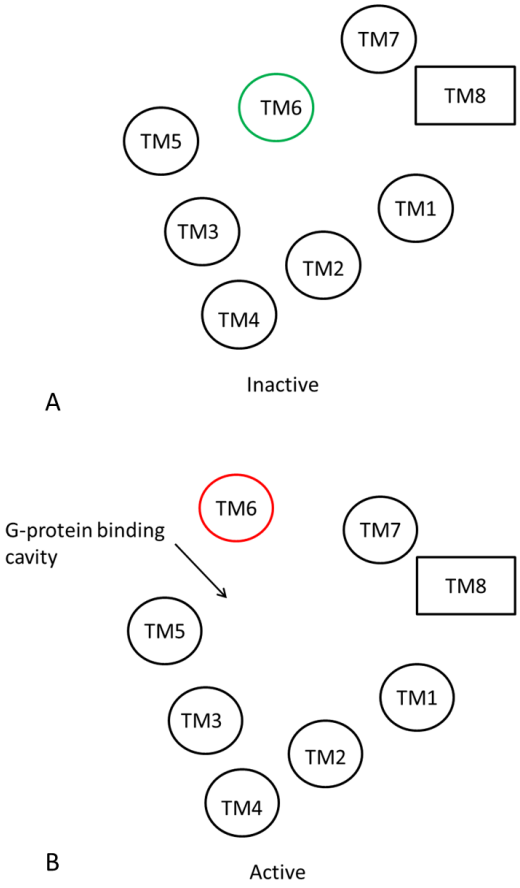


Figure 1.25. Schematic representation of the intracellular view of a G-protein coupled receptor showing inactive and active states

Transmembrane helices one to seven are represented as circles and helix eight as a rectangle. A) Inactive conformation with transmembrane helix six (TM6) blocking the G-protein binding cavity and B) the active conformation with TM6 opening up the cavity for G-protein binding.

Not all receptors have an ionic lock, as described above, but recent work has revealed that polar interactions, similar to the ionic lock between TM3 and TM6, may oscillate in the basal state (Granier et al., 2012). This oscillation could explain the basal activity of unstimulated receptors, which spontaneously change from active to inactive states. Different ligands will alter (either break or create) different intracellular contacts and this will, in turn, change the basal state equilibrium between R and R* (Venkatakrishnan et al., 2013).

1.6.6 Constitutive activity of G-protein coupled receptors

Most GPCRs show some level of basal activity, but some GPCRs have mutations which lead to constitutive activity (Parnot et al., 2002). There is no clear cut definition of constitutive activity, so a receptor is considered to be constitutively active if it has an increase in basal activity relative to its wildtype counterpart (Parnot et al., 2002). In constitutively active receptors, the equilibrium between R and R* has moved, so receptors are more often in the R* state. Mutations that lead to constitutive activity have been identified in all regions of the receptor; however, many have mutations which disrupt non-covalent intramolecular interactions, making the receptor less stable (Parnot et al., 2002). Indeed, most such mutations are found in the transmembrane domains, suggesting that they are disrupting the stabilising structure of the helices. With stabilising interactions lost, the receptor has greater flexibility, is released from the constraints of the inactive state and is no longer silenced in the absence of its agonist. The change in intramolecular interactions could reduce the energy barrier between the R and R* conformation and so move the equilibrium towards R*. Conversely, mutations leading to constitutive activation could be creating new intramolecular interactions giving a more rigid structure, leading to a receptor stabilised in the R* conformation. It seems likely that some mutations leading to constitutive activation mimic

the active conformation stabilised by agonists of the wildtype receptor. However, it also seems likely that there is more than one conformation that can activate a G-protein. The receptor may have a new and completely different active conformation that binds to the intracellular effector in a different and possibly more effective way (Venkatakrishnan et al., 2013). Mutations may also lead to receptors which are hyperactive, meaning that they require lower concentrations of ligand to reach the active R* state compared to their wildtype counterparts (see figure 1.22) (Venkatakrishnan et al., 2013).

1.6.7 G-protein coupled receptor ligands and binding sites

Although the overall architecture of GPCRs is highly conserved, particularly in the transmembrane domains, there are many differences in the length and composition of extracellular loops and N-terminal domains. These differences can affect the ligand binding, where some receptors have accessible, and others have inaccessible pockets where the ECLs have blocked access. There is a great diversity of ligand binding sites, where some receptors have small compact pockets and others have large pockets capable of accommodating large ligands, as shown in figure 1.26.

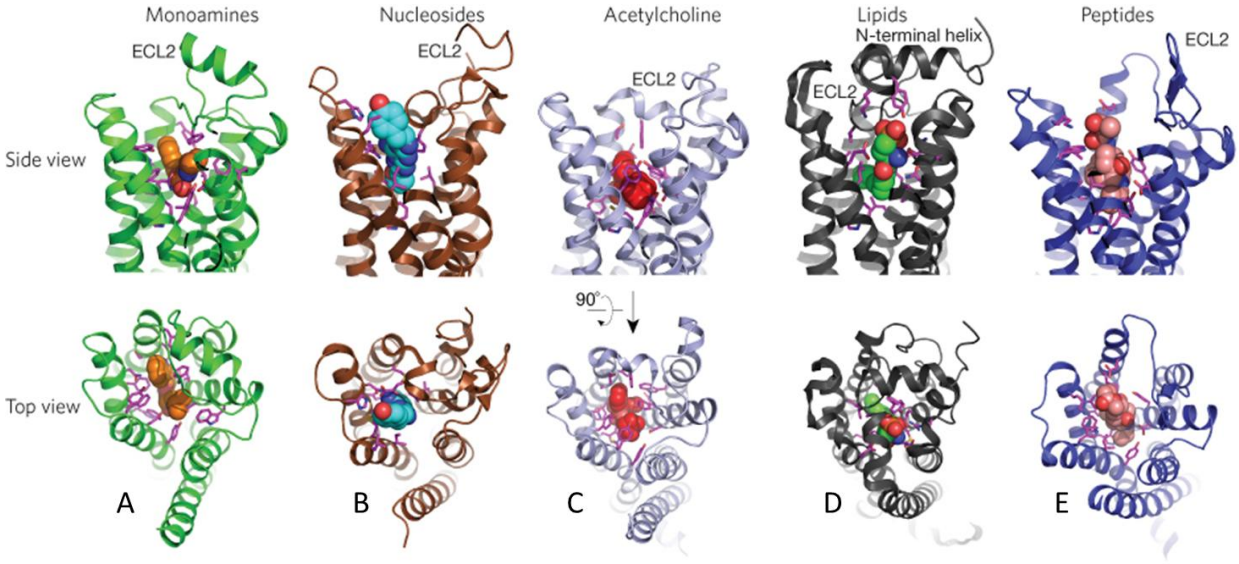


Figure 1.26. Diversity of ligand binding domains in class A G-protein coupled receptors

Side views and extracellular views of A) the β2-adrenergic receptor bound to carazolol (orange spheres), B) the A2A adenosine receptor bound to ZM24138 (blue spheres), C) the M2 muscarinic receptor bound to QNB (red spheres), D) the sphingosine 1-phosphate receptor bound to ML056 (green spheres) and E) the μ-opioid receptor bound to β-FNA (salmon spheres). Figure adapted from Granier and Kobilka (2012).

These differences indicate that receptor binding sites are specifically adapted to their ligands (Audet and Bouvier, 2012). There is also a great diversity of binding modes between receptors and ligands, which suggests that there is no unique activation trigger which is used by all GPCRs. Interestingly, G-protein coupling on the intracellular side of the receptor consistently increases the affinity of agonists and full activation of the receptor requires both agonist and G-protein engagement (Audet and Bouvier, 2012). Indeed, it has been shown that the A2AR can have an activated TM7 structure at the same time as an inactive TM6 structure. This illustrates the conformational flexibility of the receptors and also the intermediate stages of activation (Galandrin, Oligny-Longpré and Bouvier, 2007). The intermediate stage of activation is denoted R' and the states are illustrated in figure 1.27.

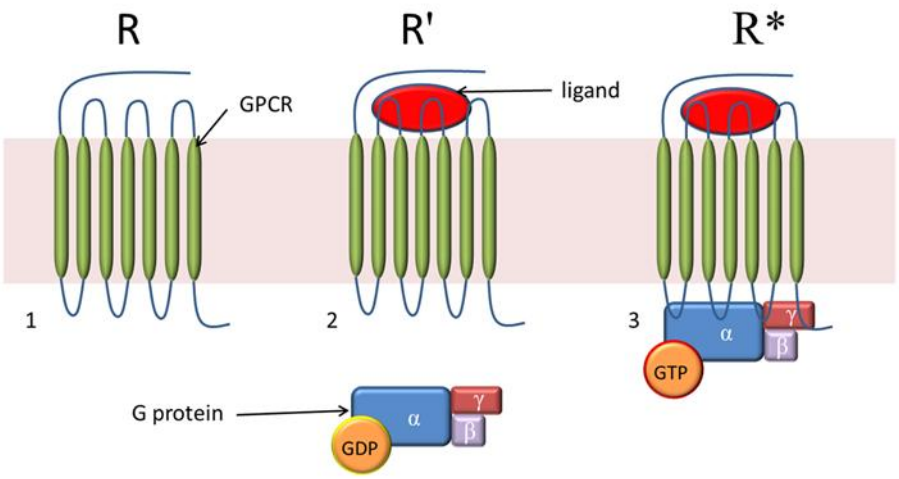


Figure 1.27. Detail of G-protein coupled receptor activation showing the inactive R, intermediate R' and active R* states

1) G-protein coupled receptor (GPCR) in the inactive R state. 2) GPCR in the intermediate R' state, bound to an agonist on the extracellular side of the receptor. 3) GPCR in the active R state with both an agonist and G-protein bound.*

Another aspect of ligand binding is a phenomenon known as biased signalling.

GPCRs primarily couple to G-proteins; however, there are other intracellular proteins which can also interact with GPCRs and can influence signalling cascades. Some ligands can preferentially trigger different intracellular signalling and this is known as biased signalling. It seems that different ligands can stabilise different conformations and this difference may affect which intracellular effector is recruited (Rahmeh et al., 2012).

1.7 Specific Features of Melanocortin Receptors

1.7.1 The melanocortin receptors

The MC1R under consideration in this study is part of the melanocortin system which comprises various peptides derived from the *POMC* gene which are thought to have co-evolved with five melanocortin receptors (MCR), (MC1R-MC5R) approximately 500 million years ago (Dores and Baron, 2011). The melanocortin receptors are involved in a diverse range of physiological functions including pigmentation (MC1R), steroidogenesis (MC2R), energy homeostasis (MC3R), regulation of food intake and sexual function (MC4R) and sebaceous gland secretions (MC5R) (Cone et al., 1993; Yang, 2011). These MCRs are all class A rhodopsin-like receptors of the superfamily of GPCRs (Ringholm et al., 2004). Figure 1.28 shows the evolutionary relationships that are thought to exist between the different GPCRs of this family, based on sequence similarity.

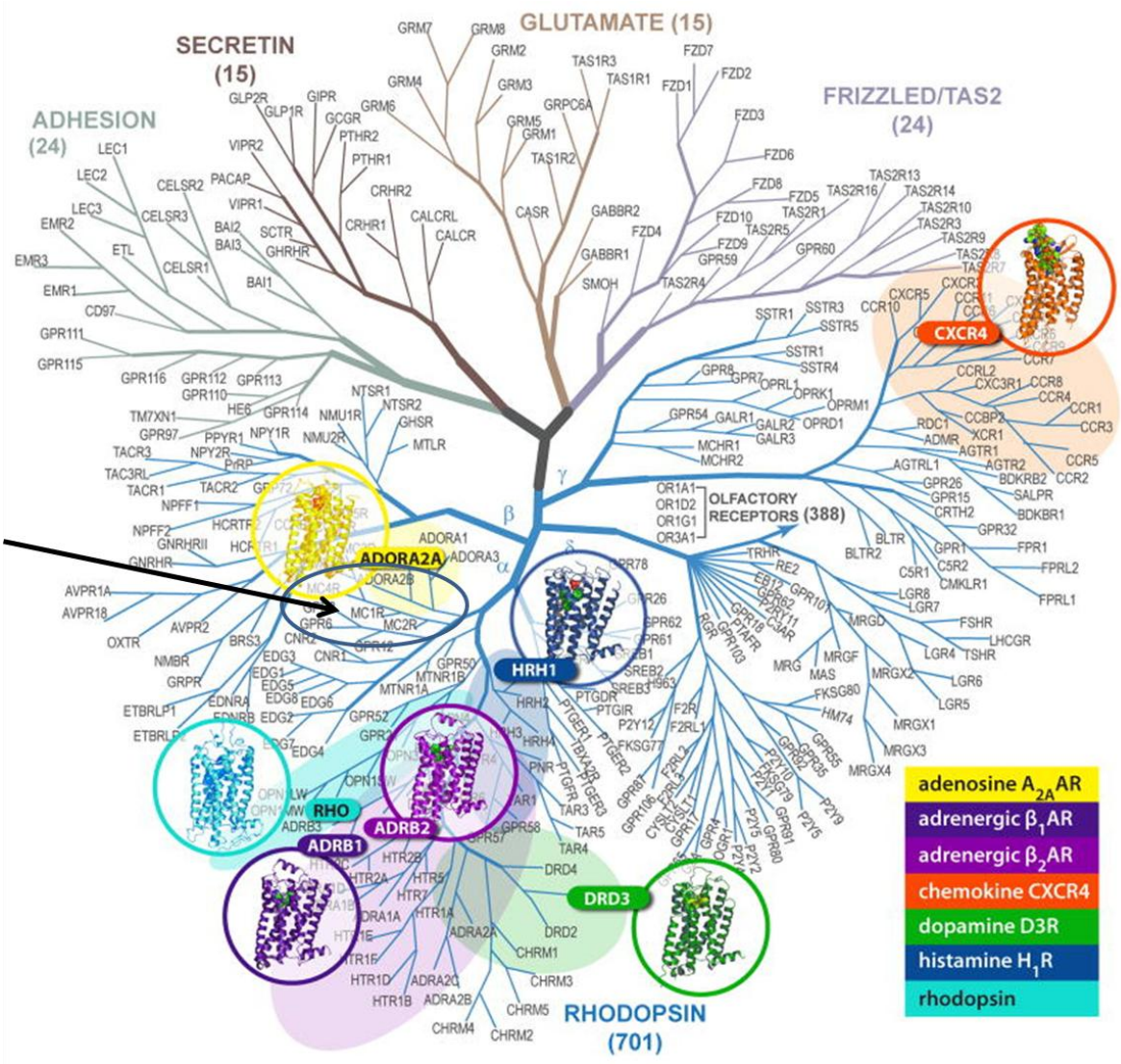


Figure 1.28. G-protein coupled receptor super family evolutionary tree

G-protein coupled receptors which have published high-resolution crystal structures are shown on the sequence homology tree. The MC1R is highlighted with the black arrow. Highlighted areas close to the structures show close homologues of the crystal structures with greater than 35% sequence identity in the transmembrane domains. Figure adapted from Katritch et al. (2012).

MCRs share many features with all GPCRs but also differ in important ways. MCRs have a high sequence similarity within the group but a comparatively low similarity to other GPCRs, the most conserved regions being in the transmembrane domains. The N-terminus of the

MC1R does not appear to be a signal peptide which is cleaved but evidence suggests that TM1 is a non-cleavable signal anchor directing the protein to its position in the membrane (Wallin and von Heijne, 1995). The N-terminal contains two glycosylation sites (see figure 1.29), N15 and N29, (human numbering) which are both glycosylated with different glycan chains in each case (Herraiz et al., 2011). This glycosylation does not affect ligand binding but is required for normal levels of cell surface expression (Herraiz et al., 2011). It is thought by Herraiz et al., (2011) that such N-glycosylation both improves forward trafficking and also decreases internalisation of the MC1R. MCRs have an unusually high degree of constitutive activity compared to other GPCRs (Holst and Schwartz, 2003). This high constitutive activity is likely to have many structural causes which create a flexible receptor: the proline in TM5, which is highly conserved in most GPCRs, is not found in MCRs. The proline is likely to alter TM packing and overall architecture of the receptor. There are thought to be two disulphide bridges between C35 and C275 and between C267 and C273 (human numbering) (Chai et al., 2005). Mutations in C267 and C275 abolish function (Frändberg et al., 2001; Holst, Elling and Schwartz, 2002) and C35 and C273 are highly conserved in all MCRs suggesting an important role in MCR function. A disulphide bridge, which is highly conserved in most GPCRs, between extracellular loop two (ECL2) and TM3, is not, however, found in MCRs (Holst and Schwartz, 2003). A diagram of a general MC1R is shown in figure 1.29.

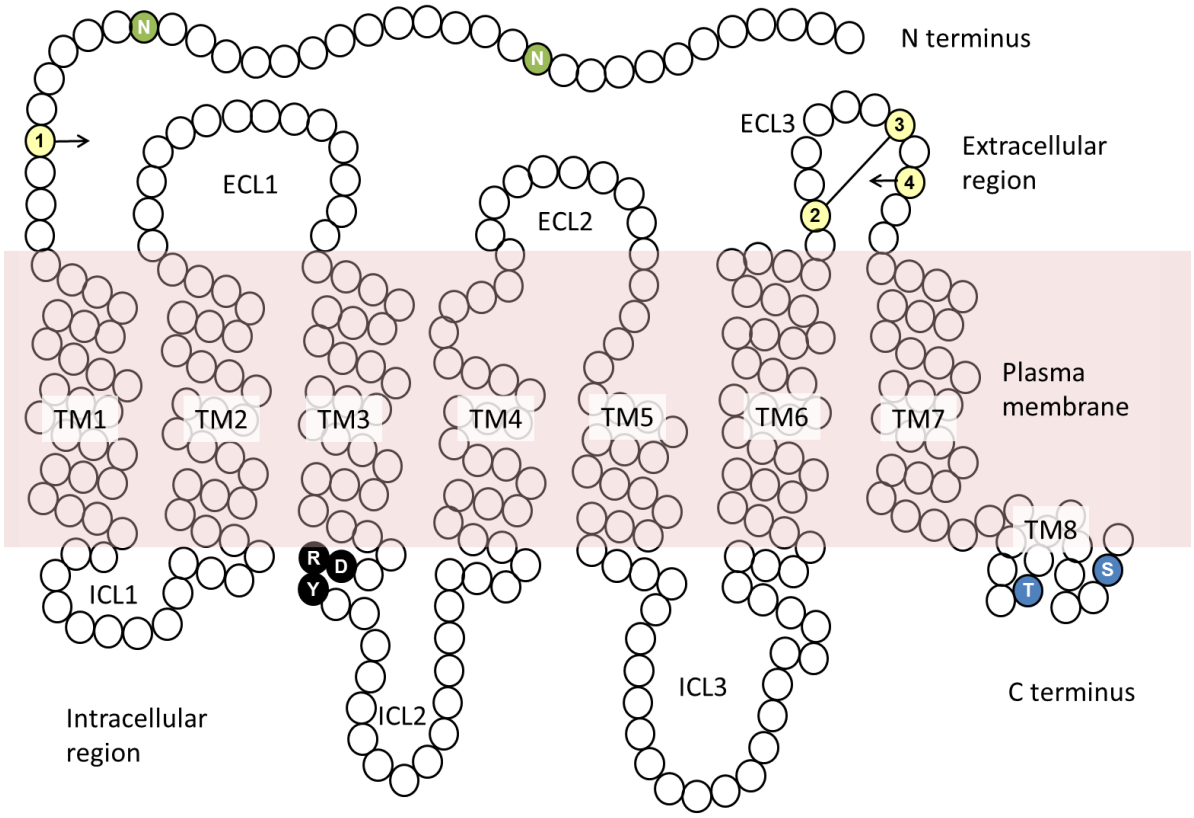


Figure 1.29. Schematic representation of the melanocortin-1 receptor

Amino acids are represented as circles. The highly conserved DRY motif is highlighted with black circles and white lettering. Asparagines of the N terminus involved in glycosylation are coloured green with white lettering. Cysteines involved with disulphide bridges are coloured yellow with numbers where 1 forms a bridge with 4 and 2 forms a bridge with 3 (Chai et al., 2003). The threonine and serine involved in phosphorylation are coloured blue (Sanchez-Laorden, Jimenez-Cervantes and Garcia-Borron, 2007). TM= transmembrane domain, ECL= extracellular loop, ICL= intracellular loop.

ECL2 is unusually short in MCRs and this is thought to contribute to their greater flexibility. The shortened ECL2 leaves the ligand binding crevice accessible for ligands. The ECL3 of MCRs is highly conserved compared to other GPCRs, particularly the Pro and Cys residues, and Chai et al., (2005) hypothesise that this gives the loop a specialised structure with a

specific function. Indeed, a point mutation in ECL3 (M281F) has been shown to be important in ligand binding (Chai et al., 2005).

1.7.2 Specific features of the melanocortin-1 receptor

Intracellular loops of GPCRs interact with heterotrimeric G-proteins, and also contain phosphorylation targets involved in internalization and recycling of the receptors. A number of naturally occurring mutations are located in intracellular loops (ICL1) and ICL2 including the "*tobacco"* mouse *(Mus musculus)* mutation S69I which leads to hyperactivity and a dark phenotype (Robbins et al., 1993). MC1Rs, like all class A GPCR receptors, have the highly conserved DRY motif at the interface between TM3 and ICL2. This DRY motif is critical for function and a mutation changing the Arg from this motif, the R142H human variant, is associated with loss-of-function and red hair. The RY of the DRY motif and the following two residues (RYIS in mouse MC1R and RYHS in human MC1R) are thought to be involved in phosphorylation by GPCR kinases (GRK) (Chai et al., 2003; Holst and Schwartz, 2003). The C-terminal extension of MC1Rs is short compared to other GPCRs. The domain is functionally relevant and is involved in the interaction of the receptor–ligand complex with G-proteins (Strader et al., 1994), correct positioning in the plasma membrane (Qanbar and Bouvier, 2003), signalling for intracellular trafficking (Schülein et al., 1998) and desensitisation and internalization of the receptor (Pitcher, Freedman and Lefkowitz, 1998; Luttrell and Lefkowitz, 2002). Truncation of the C-terminal domain Cys-Ser-Trp, "CSW" motif abolishes receptor function due to decreased expression of the receptor in the plasma membrane, and the Leu-Thr-Cys-Ser-Trp, LTCSW terminal sequence is essential for cell surface expression (Sanchez-Mas et al., 2005). A C315A mutation impairs receptor function in transfected HEK293 cells and it is thought that this is due to a lack of acylation of the

protein (Frändberg et al., 2001; Sánchez-Más et al., 2004). It is thought that this may be important for receptor trafficking or stabilization of the receptor in the membrane (Qanbar and Bouvier, 2003).

MC1Rs form dimers early during biogenesis in the endoplasmic reticulum (ER) (Sánchez-Laorden et al., 2006). These dimers can be both hetero- and homodimers and mutant forms can have a range of effects including dominant negative effects and effects on ligand binding affinity (García-Borrón, Sánchez-Laorden and Jiménez-Cervantes, 2005).

1.7.3 Life cycle of the melanocortin-1 receptor

GPCRs typically undergo desensitisation following agonist exposure (Luttrell and Lefkowitz, 2002). This desensitisation is due to phosphorylation of agonist-occupied receptors and these receptors no longer couple to their signalling pathways. Such phosphorylation is carried out by GRKs and is followed by the recruitment of β-arrestin and the formation of clathrin-coated pits, ultimately leading to endocytosis of the receptor (Drake, Shenoy and Lefkowitz, 2006). Such desensitisation has been reported for MC1Rs in heterologous expression systems (Sanchez-Mas et al., 2005). Phosphorylation is usually followed by internalization of the receptor-agonist complex. The T308 and S316 residues in humans have been shown to be important targets for phosphorylation, as shown in figure 1.29 (Sanchez-Laorden, Jimenez-Cervantes and Garcia-Borron, 2007). Internalised receptors are either degraded or recycled and other signalling events may be triggered (Swope et al., 2012). The general life cycle of an MC1R is shown in figure 1.30.

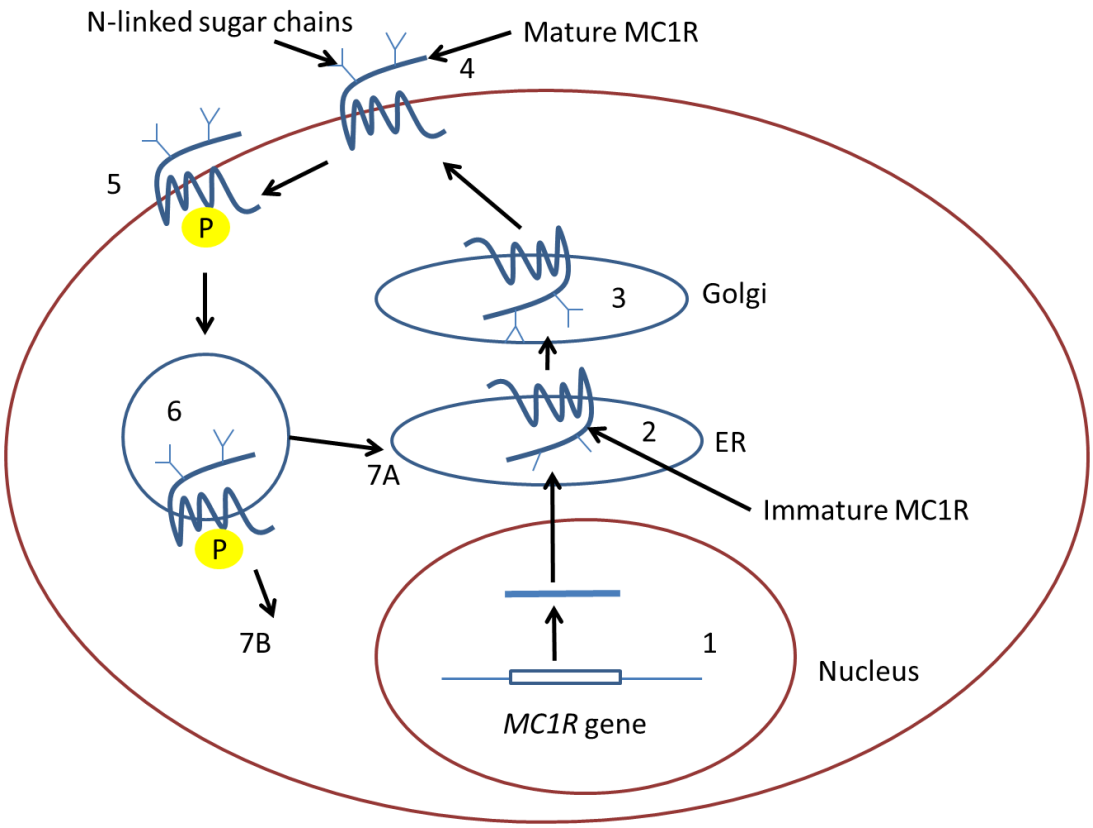


Figure 1.30. Overview of the life cycle of the melanocortin-1 receptor from transcription to maturation

1) *Transcription of the melanocortin-1 receptor (MC1R) gene in the nucleus, 2) translation of the MC1R in the endoplasmic reticulum (ER) where glycosylation is initiated, 3) further glycosylation of the MC1R in the Golgi apparatus, 4) mature, glycosylated MC1R in the plasma membrane, 5) phosphorylation of the MC1R, 6) internalisation of the MC1R followed by either 7A) recycling or 7B) degradation (Drake, Shenoy and Lefkowitz, 2006; Sanchez-Laorden, Jimenez-Cervantes and Garcia-Borron, 2007; Herraiz et al., 2011).*

1.7.4 Constitutive activity of melanocortin-1 receptors

Like other GPCRs, activation of MCRs involves inward movement of the extracellular end of TM6 towards TM3. Studies have shown that the spatial gap between TM3, TM6 and TM7 closes on activation and this is interpreted as an inward movement (Holst and Schwartz,

2003). MCRs have an unusually high level of constitutive activity and Holst et al. (2003) hypothesise that this is physiologically relevant. The unusually short ECL2 could be allowing movement of TM3 giving MCR receptors their high constitutive activity (Holst and Schwartz, 2003). Holst et al., (2003) suggest that TM6 is held towards TM3 by ECL3 which is shortened by the disulphide bridges. The ligand binding crevice of MCRs also allows free movement of TM6 in towards TM3. Agonists will stabilise this position, whereas inverse agonists would function by holding TM6 and TM7 away from TM3 (Holst and Schwartz, 2003). Like other GPCRs, MCRs have a ligand binding pocket located below the plasma membrane to which several TMs contribute (Haskell-Luevano et al., 1996). There are three residues in particular, located deep in the pocket, which are important for ligand binding and which are highly conserved in all MCRs. These residues are E92, D117 and D121 (mouse numbering) and they contribute to a highly charged, acidic region thought to be involved in ligand binding (Haskell-Luevano et al., 1996; Lu, Vage and Cone, 1998).

1.7.5 Agonists of the melanocortin-1 receptor

MCRs are unusual in having both endogenous agonists and inverse agonists. The naturally occurring agonists for MCRs are alpha, beta and gamma melanocyte-stimulating hormone (α-, β- and γ-MSH) and adrenocorticotropin (ACTH) derived from post-translational modification of the *POMC* gene transcript (Lerner, 1993) (see figure 1.31).

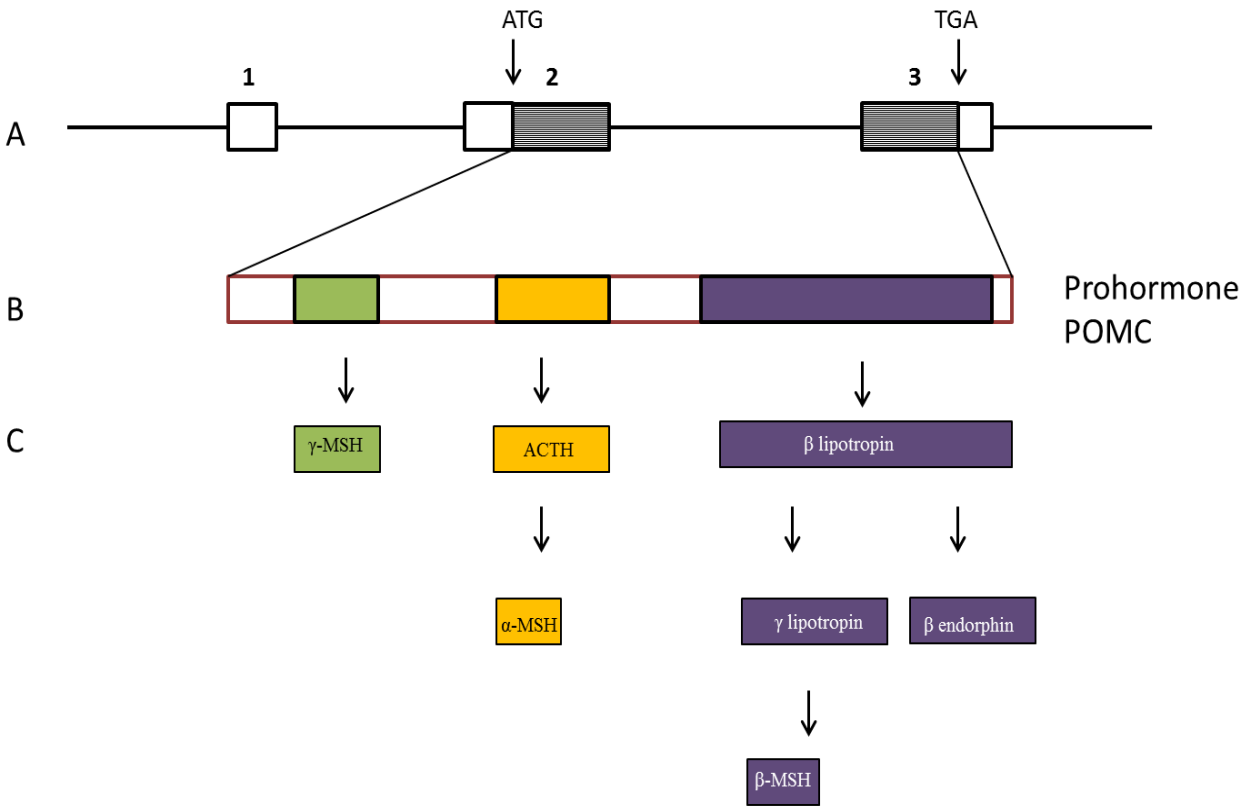


Figure 1.31. Genomic organisation of the *pro-opiomelanocortin* gene and its post-translational modifications

A) Exons of the pro-opiomelanocortin (POMC) gene are shown as boxes numbered one to three, showing the start and stop codons as vertical lines. The protein coding regions of exons two and three are indicated as boxes with vertical lines. B) The prohormone POMC indicating the different regions of the peptide which are cleaved to create the different peptides. C) The final peptide products of the POMC are shown as boxes. MSH = melanocortin stimulating hormone, ACTH = adrenocorticotropin (Lerner, 1993) .

These peptide products of the *POMC* gene all contain the highly conserved HFRW motif. The α-MSH is a 13 amino acid peptide where the HFRW are residues six to nine. Mutational and truncation studies have shown these four residues to be the minimal essential sequence required for biological activity (Haskell-Luevano et al., 1996). For agonists to function

effectively, they must be able to firstly recognise the receptor, secondly bind to the receptor and thirdly stimulate the receptor. The His (residue 6) of the HFRW motif seems to be important for stabilising the ligand-receptor complex and the aromatic ring permits conformational change to take place which is required for signal transduction (Holder and Haskell-Luevano, 2004). Phe (residue 7) also contributes to a stable complex. Holder et al., (2004) suggest that a hydrophobic aromatic network of Phe residues in the ligand and receptor interact. Arg (residue 8) is essential for activation and it is postulated that the E92, D117 and D121 (mouse numbering) form an acidic pocket allowing an ionic interaction with the Arg of the ligand. Arg also seems to be important for molecular recognition (Yang et al., 1997; Holder and Haskell-Luevano, 2004). Mutational studies have also shown that the aromatic nature of the Trp (residue 9) plays an essential role in the receptor-ligand interaction (Haskell-Luevano et al., 1996). Further studies on MC1R agonists have shown that non-peptide compounds can activate the receptor. Activation in these cases does not require the Arg (8) residue and the data suggest that a network of aromatic residues in TM4, 5 and 6, forming hydrophobic interactions during agonist binding, may be more important than the presence of an arginine-like residue in the agonist (Holder and Haskell-Luevano, 2004). Interestingly, MC1R is also modulated by zinc ions. Zinc ions alone can act as a partial agonist and at higher concentrations potentiate the action of α-MSH. It is thought that zinc ions hold ECL3 close to TM3 (Holst and Schwartz, 2003).

In melanocytes, the MC1R has a high basal level of ligand-independent constitutive activity but the receptor is activated further by α-MSH (Mountjoy et al., 1992; Sánchez-Más et al., 2004). α-MSH is secreted by the pituitary gland where it is cleaved from the precursor, pro-opiomelanocortin (POMC) (see figure 1.31). α-MSH is also produced by keratinocytes and

melanocytes in the epidermis (Chakraborty et al., 1996; Slominski et al., 2000; Walker and Gunn, 2010). The actions of α-MSH were first demonstrated about a century ago when tadpoles, incubated in pituitary extracts, turned black (Lerner and McGuire, 1961; García-Borrón, Sánchez-Laorden and Jiménez-Cervantes, 2005). It was also shown that purified α-MSH induces skin darkening in humans (Lerner, 1993). α-MSH is a small peptide (13 residues) in contrast to the agouti signalling protein (ASIP) (131 residues) which functions as an inverse agonist and is much larger and structurally very different.

1.7.6 Inverse agonists of the melanocortin-1 receptor

The endogenous inverse agonist of the MC1R is ASIP, which is encoded by the *agouti* locus. ASIP is a 131 (mouse number) residue peptide produced by dermal papillae cells where it acts in a paracrine fashion in the microenvironment of follicular melanocytes. In humans and mice, ASIP acts as a high affinity, inverse agonist of the MC1R, inhibiting binding of the agonist α-MSH, and stabilising the inactive conformation of the receptor which reduces signalling (Bultman, Michaud and Woychik, 1992; Miller et al., 1993; Miller et al., 1993; Lu et al., 1994). The exact mechanism by which signalling is reduced is not completely understood but ASIP binding to the MC1R is accompanied by reduced levels of intracellular cAMP and a switch to phaeomelanogenesis. This change is not the only effect that ASIP has on melanocytes. ASIP leads to a reduction in the rate of new MC1R biosynthesis (Rouzaud et al., 2003) as well as changes in gene expression which lead to dedifferentiation of melanocytes to melanoblasts (Le Pape et al., 2009). It has also been shown that ASIP delays terminal differentiation of ventral melanocytes in embryonic deer mice (Manceau et al.,

2011). It is not clear if this is mediated through cAMP levels or if there is another signalling pathway involved.

The *ASIP* gene consists of three protein coding exons; 2, 3 and 4, and in mice four non-coding exons; 1A, 1A', 1B and 1C, as shown in figure 1.32.

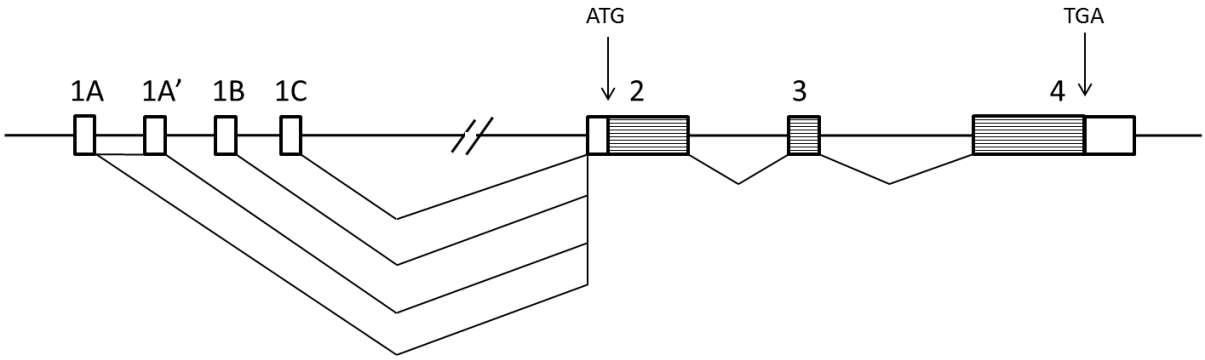


Figure 1.32. Genomic organisation of the *agouti signalling protein* gene

The broken horizontal line indicates DNA (not to scale). Exons are shown as numbered boxes. Exons 1A, 1A', 1B and 1C are untranslated exons. The regions in exons 2, 3 and 4 with horizontal stripes show protein coding sequences. Untranslated regions of exons 2, 3 and 4 are unfilled. Vertical arrows indicate the start and stop codons in exon 2 and exon 4 respectively. Diagonal lines, joining exons indicate alternative transcripts of the gene; 1A-2-3-4 and 1A-1A'-2-3-4 are ventral specific transcripts, 1B-2-3-4 and 1C-2-3-4 are dorsal, hair-cycle specific transcripts (Vrieling et al., 1994; Fontanesi et al., 2010).

This general arrangement is conserved in many other mammals, for example, rabbits, dogs, foxes, cats, mice, pigs, horses, cows and humans (Bultman, Michaud and Woychik, 1992; Kwon et al., 1994; Våge et al., 1997b; Leeb et al., 2000; Rieder et al., 2001; Eizirik et al., 2003; Kerns et al., 2004; Girardot et al., 2005; Fontanesi et al., 2010). The N-terminal 22

amino acid sequence is a cleavable signal peptide that is essential for secretion of the protein (Perry et al., 1996). In mice there are multiple regulatory elements controlling the coding sequence and it has been shown that exons 1A and 1A' are ventral specific isoforms whereas exons 1B and 1C are hair cycle specific isoforms (Vrieling et al., 1994). In the dorsal coat of wildtype *agouti* mice, a pulse of ASIP expression occurs during the mid-anagen phase of the hair cycle (days 2-7 for mice). This pulse of ASIP expression leads to a subapical band of phaeomelanin on otherwise black hairs, giving the hairs a characteristic banded appearance (see figure 1.33).

Figure 1.33. Diagrammatic representation of unbanded and banded mouse hairs

A) Unbanded, eumelanic hair and B) banded "agouti" hair showing a subapical band of phaeomelanin on an otherwise eumelanic hair. Figure adapted from Hoekstra (2006).

In contrast, the ventral coat of agouti mice is cream coloured due to the expression of ASIP throughout the anagen phase. The differences between pigmentation of the dorsum and ventrum are due to the use of different promoters for *ASIP* (Vrieling et al., 1994). The ventral specific promoter directs the continuous expression of ASIP whereas the hair cycle specific

promoter shows transient expression of ASIP (Vrieling et al., 1994). It is thought that this transient expression is related to bone morphogenetic protein (BMP) and that this protein is involved in up-regulating *ASIP* in the dermal papilla (Sharov et al., 2005).

1.7.7 Pleiotropy and the agouti signalling protein

A fascinating and widely studied example of changes to *ASIP* regulation is demonstrated in the A^{Y} allele found in mice. The allele is lethal in the homozygous state and results in a yellow phenotype in heterozygotes. The allele has a number of pleiotropic effects including obesity, non-insulin-dependent diabetes and an increased predisposition to tumours (Duhl et al., 1994; Klebig et al., 1995). The explanation for these dramatic effects of this allele is that it has a 170 kb deletion in a regulatory region, leaving *ASIP* under the control of the promoter of the housekeeping gene, *Raly,* which leads to ectopic expression of *ASIP*. A remarkably similar mutation is found in the *Y* allele of the Japenese quail which is also lethal in the homozygous state and results in a yellow phenotype in heterozygotes (Nadeau et al. 2008). Here a > 90 kb deletion upstream of *ASIP* encompasses almost the entire coding region of *Raly* and the eukaryotic initiation factor 2 gene, *EIF2B,* which plays a role in protein synthesis (Suragani et al. 2005). As in mice, *ASIP* is left under the control of the *Raly* promoter, leading to overexpression of ASIP in many tissues in the quail. This leads to pleiotropic effects including abnormal metabolism and higher levels of abdominal fat in heterozygotes (Minvielle et al. 2007). ASIP may have other effects on MC1Rs which are as yet not known. This is suggested by observations that A^{Y}/A mice on a wildtype MC1R background are paler in colour than the same mice on a non-functional MC1R background. Indeed, it has been shown that the effects of ASIP on melanocyte morphology are mediated,

as expected, through MC1R, but are unexpectedly cAMP independent (Hida et al., 2009). ASIP is only expressed in the skin of mice, but in rabbits, cattle, pigs and humans, ASIP has been found in a large range of tissues including the heart, ovaries, testes, foreskin and adipose tissue (Girardot et al., 2005; Norris and Whan, 2008; Fontanesi et al., 2010). This wide ranging expression suggests that ASIP has wider effects that are not currently understood.

1.7.8 Structure of the agouti signalling protein

Nuclear Magnetic Resonance (NMR) analysis has revealed that ten conserved cysteine residues of ASIP form a scaffold of five disulphide bonds. ASIP structure includes an unusual inhibitor cysteine knot fold motif. This fold forms a constrained and well-structured loop, created by three disulphide bonds, which functions as the active loop (McNulty et al., 2005). ASIP has three functional domains which are all essential for functional interactions with the MC1R: the C-terminal loop, the active loop and the N-terminal loop. The active loop contains the highly conserved Arg-Phe-Phe, "RFF" sequence motif which makes direct contact with the receptor in the ligand binding pocket (figure 1.34). The site of contact is thought to be partially overlapping the site for α-MSH binding (Haskell-Luevano et al., 2001; Jackson et al., 2005). The RFF motif is essential in recognition, binding and inverse agonist function of ASIP (Tota et al., 1999). The action of the RFF motif is thought to mimic the HFRW motif of α-MSH, indeed, the partial sequence CRFFNAFC of ASIP functions as an inverse agonist rather than as an agonist (Haskell–Luevano et al., 2000). All three loops of ASIP are required for full function as an inverse agonist but studies have shown that chimeras lacking the C-terminal loop have reduced binding affinities and, most remarkably, functioned as agonists (Patel et al., 2010). This suggests that the C-terminal loop is essential for function

as an inverse agonist. It has been suggested that the C-terminal loop forms a contact point with ECL1 of the MC1R; specifically the LVARAA (114-119 human numbering) sequence of MC1R and VLSLN (106-111 human numbering) of ASIP form hydrophobic interactions, as shown in figure 1.34 (Patel et al., 2010). It has been suggested that with the removal of stabilising interactions between the C-loop and ECL1, the RFF motif of the active domain relocates to the same position as HFRW motif of α-MSH and thus activates the receptor in a similar way to α-MSH (Jackson et al., 2005). The N-terminal loop of ASIP interacts with ECL2 and ECL3 of MC1R and is essential for binding (Patel et al., 2010).

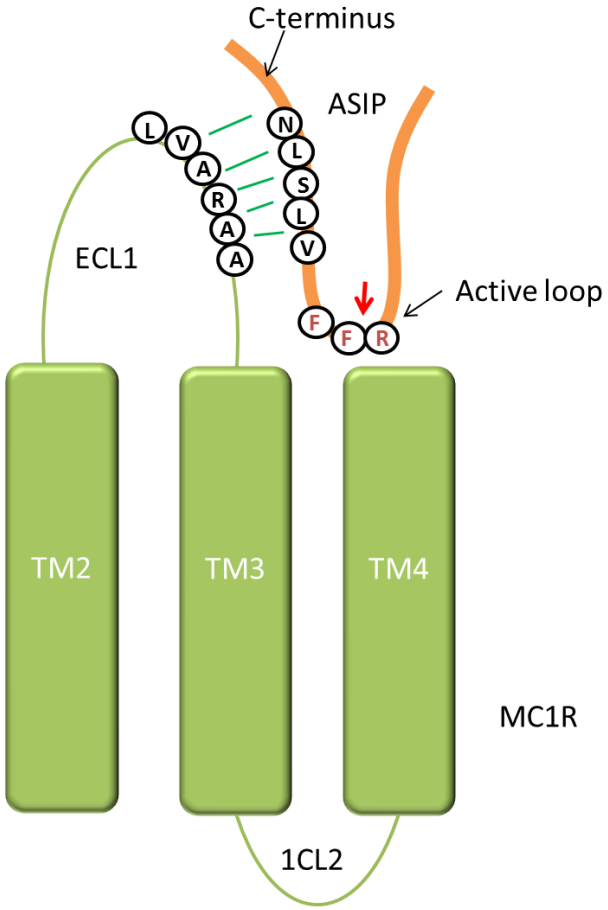


Figure 1.34. Schematic representation of the interaction between the melanocortin-1 receptor and the agouti signalling protein

A fragment of the melanocortin-1 receptor (MC1R) is shown in green including transmembrane helix two (TM2), TM3, TM4, extracellular loop one (ECL1) and intracellular loop 2 (ICL2). Hydrophobic amino acids of the human MC1R in ECL1 are shown as circles. A fragment of the agouti signalling protein (ASIP) is shown in orange including the C-terminus and the active loop. Hydrophobic amino acids of the human ASIP of the C-terminus are shown as circles with black letters (Patel et al., 2010). Hydrophobic interactions are indicated as green lines. The highly conserved RFF motif of the active loop of ASIP is shown as circles with red lettering and the red arrow indicates the movement of ASIP into the binding pocket of the MC1R..

The two state model of receptor binding predicts that agonists will preferentially bind to the R* state and that inverse agonists will preferentially bind to the R. Competitive binding

assays have demonstrated that this is not the case with MC4R and its ligands. In this case results were consistent with an induced fit model rather than the two state model (Chai et al., 2003).

1.8 Predicting Protein Structure

There are, as outlined above, many cases where mutations in the MC1R lead to differences in function of the protein. In order to understand the effects that mutations have on a protein's function, it is necessary to understand the protein's tertiary structure. Experimental methods of protein structure determination include X-ray crystallography and NMR but these methods are often unachievable due to technical, time and financial constraints. This is particularly the case, as has been mentioned, for GPCRs due to their inherent flexibility. To put the difficulty of solving GPCR structure into perspective, the Nobel Prize for Chemistry in 2012 was awarded for studies of G-protein coupled receptors, following work elucidating the structure and function of these molecules. In contrast, computational methods for protein prediction are automated, cheap and fast. However, these computational methods have limitations and prediction of protein structure remains a major unsolved problem in structural molecular biology. There is a widening gap between the number of known protein sequences, which is over 100 million, and the number of solved protein structures, which is approximately 100 000 (RCSB, 2014). Predicting a protein's structure from its amino acid sequence is a notoriously difficult problem as there are many conformations possible, due to the rotation of the chain about carbon atoms in the protein's backbone. The main difficulties facing protein prediction are calculating the free energy and finding the kinetically accessible global minimum of this energy. At present, for a computer to assess each possible conformation would take many years of computation and may still give an inaccurate model compared to the native protein (Branden and Tooze, 1991).

1.8.1 Homology modelling

Homology modelling is one of the most commonly used and effective solutions to predicting protein structure. Homology modelling uses a combination of amino acid sequence alignments and experimentally determined three-dimensional protein structures to identify homologous proteins which may be used as templates for predicting protein structure. The method relies on the observation that protein structures are more conserved than protein sequences. Indeed, it has been suggested that the number of tertiary structural motifs of proteins is limited (Baker and Sali, 2001; Kaczanowski and Zielenkiewicz, 2010). The first and most critical step in protein modelling is an amino acid alignment where a powerful algorithm is used to search for similarities between the sequence of interest and those in a library of known protein structures. These libraries may include the Structural Classification of Proteins (SCOP) database and the Protein Data Bank (PDB). One of the most powerful sorts of alignment is a position-specific iterated basic local alignment search tool (PSI-BLAST). A PSI-BLAST repeatedly searches the database using multiple alignments to generate new position-specific scoring matrices which are then used for the next round of searching. The PSI-BLAST will repeat iterations until no new sequences are found (Altschul et al., 1997). This method allows sequences to be matched to remote homologues. Once a suitable template of known structure has been identified, a three dimensional model of the protein of interest is generated and a pdb file is created. Pdb files contain the coordinates of the atoms in a molecule and these files can be opened with a range of computer programmes which allow the protein to be visualised, manipulated and/or subjected to further analysis. Any further analysis of a pdb file is only as good as the original and there are limitations to the accuracy and usefulness of such models.

1.8.2 Computer programmes for analysing molecular structures and interactions

There are a number of computer programmes, which can analyse pdb files that are used in this study. One analysis uses a programme called Socket to analyse the interactions between the seven helices within a GPCR. This analysis investigates "knobs into holes" interactions which were first proposed by Francis Crick in 1953. Crick analysed the formation of coiled coils, where two α-helices wrap around each other to form a super-helical structure. The structure is maintained by an interface of interlocking hydrophobic amino acids. Crick was able to show that a seven amino acid repeat would result in a hydrophobic stripe along the helix. The formula he gave was $(abcdef)_n = (HxxHxxx)_n$ where H denotes a hydrophobic residue. Crick proposed a "knobs into holes" model where hydrophobic amino acid knobs would become interlocked into hydrophobic holes on the partner helix (Crick, 1953). GPCRs present a complex grouping of helices and it is predicted that most of the intramolecular interactions are towards the intracellular side of the molecule (Venkatakrishnan et al., 2013). The Socket programme is able to predict the interactions between amino acids of the helices and identify the residues that act as "knobs" and those that act as "holes" (Walshaw and Woolfson, 2001). This information is useful in understanding the tertiary structures of GPCRs.

Many other programmes are available for analysing pdb files, for example "Fpocket" which is used for predicting binding sites in proteins (Le Guilloux, Schmidtke and Tuffery, 2009). Fpocket is able to identify cavities on the surface of proteins which may be potential binding pockets for ligands. Furthermore, the programme predicts the specific residues within the pocket which are likely to interact with a ligand. Taken alone, the information provided by the programme may be limited. However, taken together with information on residue

conservation from amino acid alignments, and information from experimental evidence, Fpocket predictions can contribute to the prediction of binding sites and residues involved in protein-ligand interactions.

1.8.3 Predicting molecular interactions

Many computer programmes have been developed in recent years which use pdb files to predict interactions between two molecules. Of particular relevance to this study is the interaction between the MC1R and its ligands. The process is known as "docking" and in the field of molecular modelling, docking aims to predict the orientation and strength of binding between a ligand and its receptor. Docking is a computational calculation based on a range of assumptions and approximations. The scoring systems vary from programme to programme. The programme used in the present study is *CLC Drug discovery workbench* (*CLC Discovery*) which is designed for drug discovery. Firstly, the programme identifies feasible ligand binding pockets, one of which is selected by the user. Secondly, a docking simulation is carried out where ligands are fitted into the selected binding pocket. Many random starting positions are sampled and the best results are used to iteratively improve the sampling. Docking results are scored to give an evaluation of the binding, where hydrogen bonds, ionic interactions and non-polar interactions are rewarded and non-polar to polar contacts, repulsive contacts (including hydrogen bond donor to donor contacts, hydrogen bond donor to ionic contacts and hydrogen bond acceptor to acceptor contacts) are penalised. Results may be visually inspected to check that the orientations are feasible.

There are a growing number of computer programmes and servers available for homology modelling and docking with continuous improvements to accuracy. Accuracy of homology

modelling is regularly assessed in an international blind trial of protein prediction techniques. This trial is known as Critical Assessment of Structure Prediction (CASP) and is a biennial competition aimed at advancing the field. Targets for protein prediction are proteins where the structure is soon to be solved by experimental methods, and competitors are required to predict the structure from the amino acid sequence. This competition has shown that there is a steady improvement in accuracy of prediction and shows that a sequence identity as low as 30% can be used to predict a protein structure with a root mean square deviation (RMSD) to the native structure of 2-4 Å, which is equivalent to a low resolution X ray structure (Beuming and Sherman, 2012). Unsurprisingly, the best results are obtained where there is a high sequence identity to the template structure. There are similar competitions to assess docking accuracy with GPCRs (Michino et al., 2009; Kufareva et al., 2011; Beuming and Sherman, 2012). These competitions show that models of docking to GPCRs can predict receptor-ligand complexes with a RMSD of 2.5 Å from the crystal structure. The accuracy of docking using homology models, however, is not so good, as the assumptions and approximations of the docking are compounded by the assumptions and approximations of homology models. The best predictions of docking with homology models are achieved when information from experiments (for example mutagenesis) and additional computational approaches are combined and this is the approach taken in this study (Beuming and Sherman, 2012).

1.9 Aims and Objectives

The overall aim of this study was to gain an understanding into the genetic and molecular basis of melanism in the grey, fox and red squirrel with a primary focus on the grey squirrel. There were three key areas of study: genetic, computer-based and cell-based studies.

1.9.1 The melanocortin-1 receptor

Given that variations in the *MC1R* gene are associated with melanism with a wide variety of species, as discussed, this gene was considered the first candidate to investigate the genetic and molecular basis of melanism in these three species of squirrel. This study revealed that the genetic basis of melanism was associated with a mutation in the *MC1R* gene of the grey squirrel. This finding led to a computer-based investigation into the three dimensional structural differences between the wildtype and mutant MC1R of the grey squirrel. The finding also led on to an investigation into how the mutation affects the function of the MC1R in the cell. Different mutations artificially introduced into both the wildtype and mutant MC1R of the grey squirrel were investigated in order to further understand the specific residues involved in receptor function.

1.9.2 The agouti signalling protein

A secondary aim of this study was to investigate the genetic basis of melanism in the fox and red squirrels. As no mutations in the MC1R were found to be associated with melanism in these species, the study was extended to investigate another gene. Given that the *ASIP* gene also plays a key role in pigmentation, as discussed, this gene was considered as a second candidate to investigate the genetic basis of melanism in the fox and red squirrels. Finally, the

POMC gene encoding the α MSH was also sequenced in order to verify the sequence of this MC1R agonist.

Objective 1: Sequence the *MC1R* gene of the wildtype and melanic morphs of the grey, fox and red squirrels.

Hypothesis: Variations in the *MC1R* gene are associated with melanism in the grey, fox and red squirrels.

Objective 2: Sequence the protein-coding exons of the *ASIP* gene of the wildtype and melanic morphs of the grey, fox and red squirrels.

Hypothesis: Variations in the *ASIP* gene of the fox and red squirrel are associated with melanism in the fox and red squirrel.

Objective 3: Sequence the region of the *POMC* gene encoding the α MSH of the grey squirrel.

Hypothesis: The sequence of the α-MSH in the grey squirrel is identical to that of all other mammals.

Objective 4: Investigate the structure of the MC1R of the wildtype and melanic grey squirrel using computer-based predictions.

Hypothesis: Sequence variation between the wildtype and melanic *MC1R* gene leads to differences in structure of the MC1R protein between the two morphs.

Objective 5: Investigate the interactions between the MC1R of both the wildtype and melanic grey squirrel with their ligands, α-MSH and ASIP, using computer software.
Hypothesis: Differences in MC1R structure between the wildtype and melanic morph lead to differences in receptor-ligand interactions.

Objective 6: Characterise the function of the MC1R from both the wildtype and melanic grey squirrel in a cell-based system by comparing the cAMP response of cells transfected with either the wildtype or mutant receptor. Measure levels of intracellular cAMP in cells transfected with either the wildtype or mutant MC1R in response to the ligands α-MSH and ASIP.
Hypothesis 1: Differences in the structure of the MC1R of the wildtype and melanic grey squirrel lead to the mutant receptor being constitutively active.
Hypothesis 2: Differences in the structure of the MC1R between the wildtype and melanic grey squirrel lead to the melanic MC1R being unresponsive to the inverse agonist ASIP.

Objective 7: Investigate the role of specific residues and size of deletions on the function of the wildtype and melanic MC1R of the grey squirrel.
Hypothesis 1: Deletions of two amino acids from the wildtype MC1R lead to a constitutively active receptor.
Hypothesis 2: Deletions of four amino acids from the wildtype MC1R lead to a constitutively active receptor.
Hypothesis 3: Substitution of the E91 residue with K91 leads to a constitutively active receptor in both the wildtype and melanic MC1R.

Hypothesis 4: Substitution of the N88 residue with A88 in the wildtype MC1R leads to a constitutively active receptor.

Chapter Two

Materials and Methods

2.1 Whole Squirrel Analysis

2.1.1 Squirrel samples

Most squirrel samples were from animals which had died of natural causes and which were obtained through personal contacts, named after the region of origin below. Some samples from Cambridgeshire were obtained from a pest control company; DRE Pest Control Ltd., 9 Wallmans Lane, Swavesey, Cambridge, CB24 4QY. Grey squirrel samples were obtained from Cumbria (personal contact: Jayne Thompson), West Yorkshire (personal contact: Paul Kaiserman), Northamptonshire (personal contact: Ged Howard) and Cambridgeshire (personal contacts: Dr Andrew Smith, Dr Linda King, Dr Sheila Pankhurst, and DRE Pest Control Ltd.) in Britain and from Massachusetts, Virginia (personal contact: Dr Nancy Moncrief) and British Columbia (personal contact: Professor Buzz Hoagland) in North America. Of the 51 wildtype samples, 45 were from Britain and six from North America. Of the 44 melanic samples, 36 were from Britain and eight were from North America. All melanic grey squirrel samples from Britain were obtained from Cambridgeshire, as melanic grey squirrels were not available from the other British locations tested. A total of nine fox squirrel samples were obtained from Georgia (personal contact: Dr Nancy Moncrief), North America, four being wildtype and five melanic. A total of 39 red squirrel samples were obtained, including 33 from Italy of which 10 were wildtype and 23 melanic. All Italian samples were obtained from the Friuli-Venezia Giulia region in north-eastern Italy: 30 from Udine, one from Pordenone, one from Gorizia and one from Trento (personal contact: Luca Lapini). A further six wildtype red squirrel samples were obtained from Inverness-shire and Cumbria (personal contact: Dr Sheila Pankhurst) in Britain. Tissue obtained from the

squirrels was muscle tissue from the hind leg except in the case of samples from Cambridgeshire where whole animal carcasses were obtained.

2.1.2 Phenotype analysis

Representatives of each colour morph (wildtype, brown-black and jet black) of the grey squirrel were used to describe the phenotypes. Hairs from dorsal, flank and ventral surfaces were examined and the lengths of colour bands on individual hairs were measured using a light microscope equipped with a calibrated graticule.

Note that compositions of all buffers are found in appendix 1 unless they are supplied in a kit.

2.2 DNA Analysis

2.2.1 DNA extraction

DNA extraction from squirrel tissue was carried out using a DNeasy Blood and Tissue Kit (Qiagen). Muscle tissue (25 mg) was taken from the hind leg of the squirrel and cut into small pieces and placed in a microcentrifuge tube. ATL buffer (180 µl) was added and the tissue was ground up with a sterile pestle. Proteinase K (20 µl) was added and the mixture was vortexed and incubated at 56°C for up to three hours or until lysis was complete, vortexing occasionally during incubation. The sample was vortexed for 15 seconds and 200 µl AL buffer was added to the sample and mixed thoroughly by vortexing. Ethanol (96–100%) (200 µl) was added and again mixed thoroughly by vortexing. The mixture was applied to the DNeasy mini spin column and this was centrifuged at 8000 rpm for 1 minute. Flow-through was discarded. AW1 wash buffer (500 µl) was applied to the column and this was centrifuged for 1 minute at 8000 rpm. Flow-through was discarded. This wash procedure was repeated

with 500 µl AW2 wash buffer and the membrane was dried by a further spin at 13 000 rpm for 1 minute. DNA was eluted into a microcentrifuge tube with 200 µl AE buffer directly onto the DNeasy membrane. This was incubated at room temperature for 1 minute and then centrifuged for 1 minute at 8000 rpm.

2.2.2 Primer design for the *melanocortin-1 receptor* gene

In order to amplify the *MC1R* gene, specific primers were designed using alignments of the gene from other closely related species including rat, mouse and rabbit. The sequences of conserved areas were selected and used to design primers using the web server Primer3 (Untergasser et al., 2012). Further primers were designed based on sequences obtained throughout the study. Figure 2.1 shows a map of primer positions on the *MC1R* gene.

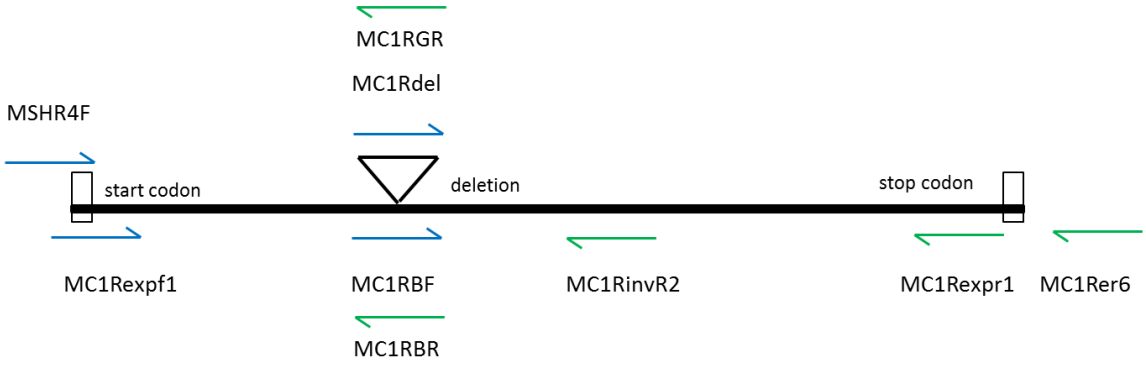


Figure 2.1. Primer map to show relative positions of primers on the *MC1R* gene

MC1Rdel matches the sequence deleted in the melanic squirrel making this a MC1R-wt specific primer and MC1RBF and MC1RBR span across the region of the deletion making these MC1RΔ24 specific primers.

2.2.3 Polymerase chain reaction (PCR) of the *melanocortin-1 receptor* gene

Genomic DNA, extracted from muscle tissue, was used to amplify the *MC1R* gene. The *MC1R* was amplified by polymerase chain reaction (PCR) in two stages. All primer sequences for PCR of genomic DNA are listed in appendix 2. The first 400 bp, including the start codon at the beginning of the gene, were amplified using the primers MSHR4F and MC1RinvR2. The rest of the *MC1R-wt* allele, including the stop codon at the 3′ end, was amplified using MC1Rdel and MC1Rer6. A forward primer, MC1RBF, and reverse primer, MC1RBR, were used to sequence the *MC1RΔ24* allele of the grey squirrel. A further reverse primer, MC1RGR, was used to sequence the *MC1R-wt* allele of the grey squirrel from heterozygotes. All PCR reactions were carried out in duplicate on a DNA thermocycler (Techne touchgene gradient) in a total volume of 25 µl using approximately 25 ng template DNA, 1 × TopTaq PCR buffer, 3 mM $MgCl_2$, 0.2 mM deoxyribonucleotide triphosphates

(dNTPs), 0.4 μM primers, and 1 × TopTaq polymerase. The following PCR conditions were used: initial denaturation 94°C for 2 minutes, 30 cycles of 94°C for 30 seconds, 59°C for 30 seconds, and 72°C for 1 minute, followed by a final extension at 72°C for 5 minutes. PCR products were purified and sent for sequencing at Source BioScience, Cambridge.

2.2.4 Primer design for the *POMC* gene

In order to amplify the *POMC* gene, human and bovine gene sequences of exon 3 from the *POMC* gene flanking the α-MSH encoding region were aligned. Conserved regions were identified and primers were designed using Primer3 (Untergasser et al., 2012). Figure 2.2 shows a map of primer positions for the *POMC* gene.

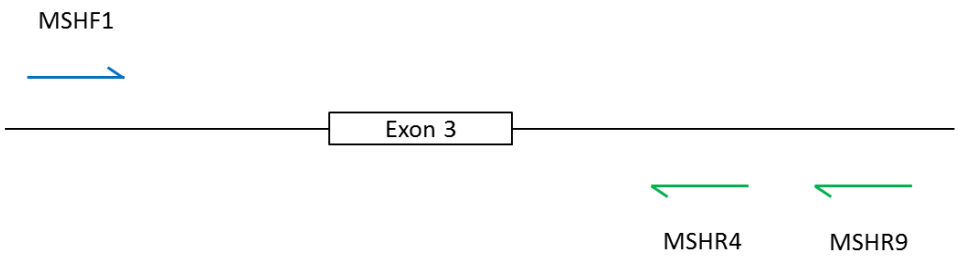


Figure 2.2. Primer map to show relative positions of the primers used to amplify the *POMC* gene where exon three contains the sequence encoding α-MSH

2.2.5 PCR of the α-melanocyte stimulating hormone gene

The sequence encoding the α-MSH for the grey squirrel was obtained by conducting two PCR reactions: the first reaction used the primers MSHF1 and MSHR9 in a total volume of 25 µl using approximately 25 ng template DNA, 1 × TopTaq PCR buffer, 3 mM $MgCl_2$, 0.2 mM dNTPs, 0.4 µM primers, and 1 × TopTaq polymerase. The following PCR conditions were used: initial denaturation 94°C for 2 minutes, 30 cycles of 94°C for 30 seconds, 67°C for 30 seconds and 72°C for 1 minute followed by a final extension at 72°C for 5 minutes. The product of this reaction was used to perform nested PCR using MSHF1 and MSHR4 using the same conditions as above but with an annealing temperature of 60 °C. This produced three bands following electrophoresis on an agarose gel and the band of approximately 600 bp was extracted by gel extraction and sequenced.

2.2.6 Primer design for the *agouti signalling protein* gene

In order to amplify the *ASIP* gene, specific primers were designed using alignments of the genes from other closely related species including the thirteen-lined squirrel, rat, mouse and rabbit. The sequences of conserved areas were selected and used to design primers using Primer3 (Untergasser et al., 2012). Further primers were designed based on sequences obtained throughout the study. Figure 2.3 shows a map of primer positions on the *ASIP* gene.

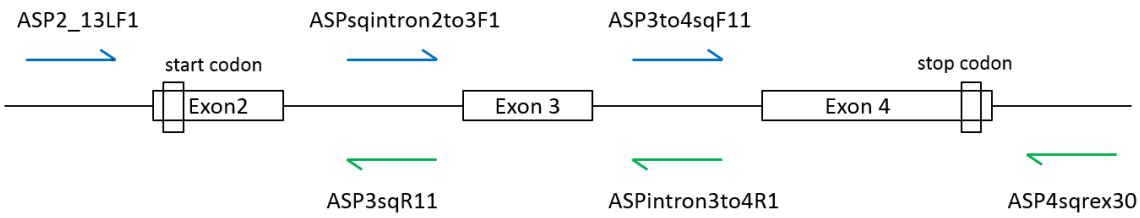


Figure 2.3. Primer map to show positions of primers on the *agouti signalling protein* gene

2.2.7 PCR for the *agouti signalling protein* gene

The *ASIP* gene was amplified as separate exons as the whole gene was too large to amplify in a single reaction. Exon 2 was amplified using the primers ASP2_13LF1 and ASP3sqR11 in a total volume of 25µl using approximately 25 ng template DNA, 1 × TopTaq PCR buffer, 3 mM $MgCl_2$, 0.2 mM dNTPs, 0.4 µM primers, and 1 × TopTaq polymerase. The following PCR conditions were used: initial denaturation 94 °C for 2 minutes, 30 cycles of 94 °C for 30 seconds, 64 °C for 30 seconds, and 72 °C for 45 seconds followed by a final extension at 72 °C for 5 minutes. Exon 3 was amplified using the primers ASPsqintron2to3F1 and ASPintron3to4R1. The PCR reaction was the same as for exon 2 except the annealing

temperature was 55ºC. Exon 4 was amplified using the primers ASP3to4sqF11 and ASP4sqrex30. The PCR reaction was the same as for exon 2 except the annealing temperature was 62 °C. All PCR products were purified and sent for sequencing at Source BioScience, Cambridge.

2.2.8 Agarose gel

Agarose (1.3 g) was added to 100 ml 1 × TBE and heated in a microwave for 3 minutes or until the agarose had dissolved. Ethidium bromide (1 µl) was added and the liquid gel was poured into gel tanks containing appropriate combs and left to set. Once set, the gel tray was loaded into the gel tank and 1 × TBE was added. The combs were removed and DNA samples were loaded with loading dye along with a DNA Hyperladder 1 (Bioline). The gel was run for approximately 1 hour at 100 V. Gels were visualised on a benchtop UV transilluminator (UVP).

2.2.9 PCR purification

Purification of PCR products were carried out using a PCR purification kit (Qiagen). Five volumes of PB buffer were added to one volume of the PCR sample and mixed. This mixture was applied to a QIAquick column and centrifuged for 1 minute. The flow-through was discarded and 0.75 ml PE wash buffer was added and the QIAquick column was centrifuged 1 minute at 13 000 rpm. The flow-through was discarded and the column was centrifuged again at 13 000 rpm to dry the membrane. DNA was eluted into a microcentrifuge tube by adding 50 µl EB buffer. This was incubated at room temperature for 1 minute and then centrifuged for 1 minute at 13 000 rpm.

2.2.10 Gel extraction

The DNA band to be extracted from agarose gel was visualised on a benchtop UV transilluminator (UVP) and excised with a clean, sharp scalpel. After excess gel had been removed, the gel slice was weighed in a microcentrifuge tube, and three volumes of QG buffer to one volume of gel (100 mg ~ 100 µl) were added. The gel was incubated at 50°C for 10 minutes and was mixed by vortexing the tube every 2–3 minutes during the incubation. When the gel slice had dissolved completely, the sample was applied to the QIAquick column, and centrifuged for 1 minute at 13000 rpm. Flow-through was discarded and 0.5ml of QG buffer was applied to the column. 0.75 ml of PE wash buffer was added to column and centrifuged for 1 minute at 13 000 rpm. Flow-through was discarded and the membrane was dried by a further spin for 1 minute at 13 000 rpm. DNA was eluted with 30 µl of EB buffer.

2.2.11 PCR for functional studies

In order to investigate the functional properties of the *MC1R* gene variants, the gene was amplified by PCR to be cloned into an expression vector. The vector used was designed to create a MC1R-Green fluorescent protein (GFP) fusion protein. Total genomic DNA was extracted from muscle tissue from wildtype and melanic grey squirrels. The entire coding sequences of the wildtype *MC1R-wt* E^{+} (accession number EU604831.2) and mutant *MC1RΔ24* E^{B} (accession number EU604830.3) genes were amplified by PCR using the primers MC1Rexpf1 (5'-CAC C**AT G**CC TGT ACA GAG GAG GCT CC-3') and MC1Rexpr1 (5'-CCC AGG AGC ACA GCA GCA CCT CC-3'). The underlined area of the forward primer highlights the Kozak sequence and the start codon is in bold. Underlined in

the reverse primer is the C added to keep the MC1R in frame with the GFP tag. This reaction was carried out in a total volume of 25 µl using approximately 25 ng template DNA, 1 × TopTaq PCR buffer, 3 mM $MgCl_2$, 0.2 mM dNTPs, 0.4 µM primers, and 1 × TopTaq polymerase. The following PCR conditions were used: initial denaturation 94°C for 2 minutes, 30 cycles of 94°C for 30 seconds, 68°C for 30 seconds, and 72°C for 1 minute followed by a final extension at 72 °C for 30 minutes. This long final extension was to ensure all products were of full length and had been 3' adenlyated. PCR products were purified and sent for sequencing at Source BioScience, Cambridge and sequences were aligned using Clustal Omega. The products were cloned into the appropriate expression vector. Plasmids were then sequenced to ensure no mutations were present.

2.2.12 Cloning of PCR products into an expression vector

The *MC1R-wt* and *MC1RΔ24* genes were cloned into a GFP Fusion TOPO TA Expression kit (Invitrogen). A recombination reaction was set up with 2 µl of fresh PCR product, 1 µl salt solution, 1 µl TOPO vector and 1 µl ddH_2O. This reaction was incubated for 5 minutes at room temperature then placed on ice. This recombination reaction was used to transform chemically competent cells.

2.2.13 Transformation of competent cells

A 50-µl vial of DH5α™-T1R cells was thawed on ice for 5-7 minutes for each transformation. Recombination product (2 µl) was pipetted gently into each vial of cells and mixed by gentle tapping. Cells were covered completely with ice, and incubated for 12 minutes. Cells were heat shocked for exactly 30 seconds in the 42°C water bath and then

replaced in ice for 2 minutes. Cells were removed from ice and 250 µl of pre-warmed super optimal broth with catabolite repression (SOC) medium was added. Cells were shaken at 37°C for exactly 1 hour at 225 rpm in a shaking incubator. Cells (100 µl) were spread onto LB agar plates containing ampicillin (50µg/ml), inverted and incubated overnight in a 37°C incubator. Single colonies were then grown up overnight in 10 ml of LB broth containing ampicillin (50µg/ml).

2.2.14 Transfection grade plasmid extraction

Plasmids for transfection were extracted from bacteria using a transfection grade mini prep kit (Qiagen). A single colony was picked from a freshly streaked selective agar plate and grown up over night with vigorous shaking at 37°C. Bacterial cells were harvested by centrifugation at 6000 rpm for 15 minutes at 4°C in a 1.5 ml microcentrifuge tube. The bacterial pellet was resuspended in 0.3 ml of P1 buffer containing RNase. P2 buffer (0.3 ml) was added and the tube mixed thoroughly by vigorously inverting. The tube was left to incubate at room temperature (15–25°C) for no more than 5 minutes. Chilled P3 buffer (0.3 ml) was added and mixed immediately and thoroughly by vigorously inverting 4–6 times. The tube was incubated on ice for 5 minutes and then centrifuged at 13 000 rpm for 10 minutes. A QIAGEN-tip 20 was equilibrated by applying 1 ml QBT buffer. The column was emptied by gravity flow. The supernatant was applied to the equilibrated QIAGEN-tip 20 and allowed to enter the resin by gravity flow. The QIAGEN-tip 20 was washed with 2 x 2 ml QC buffer. The buffer was allowed to move through the QIAGEN-tip by gravity flow. DNA was eluted with 0.8 ml QF buffer. The eluate was collected in a microcentrifuge tube and DNA was precipitated by adding 0.56 ml of room-temperature isopropanol. The tube was centrifuged immediately at 13 000 rpm for 30 minutes in a microcentrifuge at 4^{o}C. The

supernatant was carefully decanted and the DNA pellet was washed with 1 ml of 70% ethanol and centrifuged at 13 000 rpm for 10 minutes. The supernatant was carefully decanted and the pellet was air dried for 5–10 minutes and redissolved in 50 µl TE buffer.

2.2.15 Site-directed mutagenesis

In order to create the mutants used in the functional studies, plasmids containing the *MC1R* gene were subjected to site-directed mutagenesis using GENEART® Site-Directed Mutagenesis System (Invitrogen). Primers were designed with the mutation in the middle of a 35-40 bp sequence. Primers used to create each mutant are listed in appendix 2. Plasmids containing the *MC1R-wt* gene were used as template in each case, except the BGFPE91K mutant where plasmids containing the *MC1RΔ24* gene were used. For the GFPN88A mutant the primers GGFPN88AF and GGFPN88AR were used, for the GGFPE91K mutant the primers GGFPE91KF and GGFPE91KR were used, for the GGFPdel4 mutant the primers GGFPdel4F and GGFPdel4R were used, for the GGFPdel2 mutant the primers GGFPdel2F and GGFPdel2R were used and for the BGFPE91K mutant the primers BGFPE91KF and BGFPE91KR were used.

A PCR reaction was set up in a total volume of 50µl using 20 ng plasmid, 1 × enhancer, 0.3µM of each primer, 4 units of DNA methylase, 1 × SAM, 1 unit of AccuPrime™ Pfx using the following conditions: 37ºC for 20 minutes, initial denaturation 94ºC for 2 minutes followed by 18 cycles of 94ºC for 20 seconds, 57ºC for 30 seconds, 68ºC for 4 minutes (30 seconds/kb of plasmid) and a final extension of 68ºC for 5 minutes. PCR products were then recombined using the following conditions: 1 × reaction buffer, 10µl ddH_2O, 4µl PCR sample, 1 × enzyme mix. The reaction was mixed by pipetting and incubated at room

temperature for 10 minutes. The reaction was stopped by adding 1 µl 0.5 M EDTA. This product was then used for transformation of competent cells.

2.3 Cellular Analysis

2.3.1 Routine cell culture and transfection

HEK293T cells were routinely maintained in DMEM Glutamax (Invitrogen) supplemented with 10% foetal bovine serum and 100U/ml of penicillin and 100 µg/ml streptomycin sulphate in 5% CO_2 at 37ºC. Cells were seeded at 1 $\times 10^5$ cells/ml into 24 well plates, in the absence of antibiotic and grown for 24 hours until they were 80% confluent. Cells were then transiently transfected with expression vectors at 1µg/µl using Fugene (Promega) following the manufacturer's instructions with 0.5 µg DNA, 1.5 µl Fugene in 25µl Opti-MEM (Life Technologies) per well. Visualisation with a fluorescence microscope and flow cytometry analysis were used to confirm transfection and protein expression 24 hours after transfection.

2.3.2 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

To detect the presence of the MC1R-GFP fusion proteins, Western blotting was carried out on transfected cells. Firstly, SDS-PAGE was carried out on cell lysates. Ice cold RIPA buffer (200 µl) (Life Technologies) were added to each well containing transfected cells. Cells were shaken on ice for 20 minutes and then all samples were centrifuged at 13 000 rpm for 10 minutes. The supernatant was removed and 85 µl 4 × LDS sample buffer (Invitrogen) and 5µl mercaptoethanol were added. Samples were then heated on a heat block at 70˚C for 10 minutes. Each sample (5 µl) was added to each well of a pre-cast Bolt 10% bis-tris gel (Life Technologies) and 5 µl prestained protein ladder (Thermoscientific) was loaded. The gel was

run in 1 × MOPS running buffer (Life Technologies) and run at 100 V until the dye had reached the bottom of the gel (approximately 1 hour).

2.3.3 Western blotting

Following separation by SDS-PAGE, proteins were immediately transferred to nitrocellulose membranes (Biorad). Membranes and filter paper were pre-soaked in transfer buffer. The SDS-PAGE gel was carefully removed from the gel cassette and the gel and membrane were placed between filter papers and sponges. Proteins were transferred in a gel tank run at 400 mA for 1 hour. When the protein had transferred onto the membrane, the blot was incubated for 1 hour in TBS containing 0.05% (v/v) Tween 20 (TBS-Tween) with 5% BSA on a rotary shaker. This was followed by incubation in 15 ml of primary antibody at 4^{o}C overnight. Primary antibody was GFP rabbit serum polyclonal antibody (Invitrogen) at 1:10 000 dilution in TBS-Tween. Primary antibody was rinsed off with three 5 minute washes in TBS-Tween. The blot was incubated for 1 hour in 15 ml secondary antibody, goat anti-rabbit polyclonal antibody (Invitrogen) at 1:10 000 dilution in TBS-Tween. This was rinsed off with three 5 minute washes in TBS-Tween. The blot was placed in plastic wrap and incubated in 1 ml of reagent A and 1 ml of reagent B solutions from an enhanced chemiluminescent kit (Pierce) according to the manufacturer's instructions. The blot was removed from the wrap and, in a dark room, placed in a light-proof cassette with photographic paper and exposed for 30 seconds to 1 minute.

2.3.4 Fluorescence microscopy

Cells were grown, in parallel, for inspection by fluorescence microscopy, in order to directly detect the GFP of the MC1R-GFP fusion proteins. Cells were grown on coverslips in wells as described above (2.3.1) and transfected as described. Slides were prepared with a drop of mounting medium (Vectashield, Vector Laboratories). Cells were rinsed with phosphate buffered saline (PBS) and coverslips were removed from the wells and placed, cell side down, onto the mounting medium. Cells were visualised using a fluorescence microscope (Leica HC) using the appropriate filter (N21) and images were obtained using a mounted camera (Progres C5, Jenptik) and related software (Progres CapturePro 2.7).

2.3.5 Flow cytometry

Cellular transfection efficiency was assessed by flow cytometry. Three wells were assessed for all MC1R variants and these were grown in parallel wells to those grown for cAMP assays as described in 2.3.6. Cells were harvested and resuspended in 500 µl of PBS, then analysed using a flow cytometer (BD-FACSCalibur) according to the manufacturer's instructions using a 488-561 nm laser (FL1) to measure fluorescence intensity. Data were analysed using the CellQuest Pro software. Figure 2.4 shows a typical histogram of the flow cytometry data with cell counts plotted against fluorescence intensity. Cells were considered to be transfected if they fell in the "GFP positive" range, as shown in figure 2.4 (Tung et al., 2007). This histogram shows two distinct populations of cells as two peaks corresponding to (A) untransfected cells and (B) transfected cells. The areas under the curves overlap and here, cells were counted as transfected if they fell to the right of the white arrow indicated on the x axis. This inevitably excludes some transfected cells and includes some untransfected cells

but this point is considered a suitable compromise for assessing transfection efficiency (Tung et al., 2007).

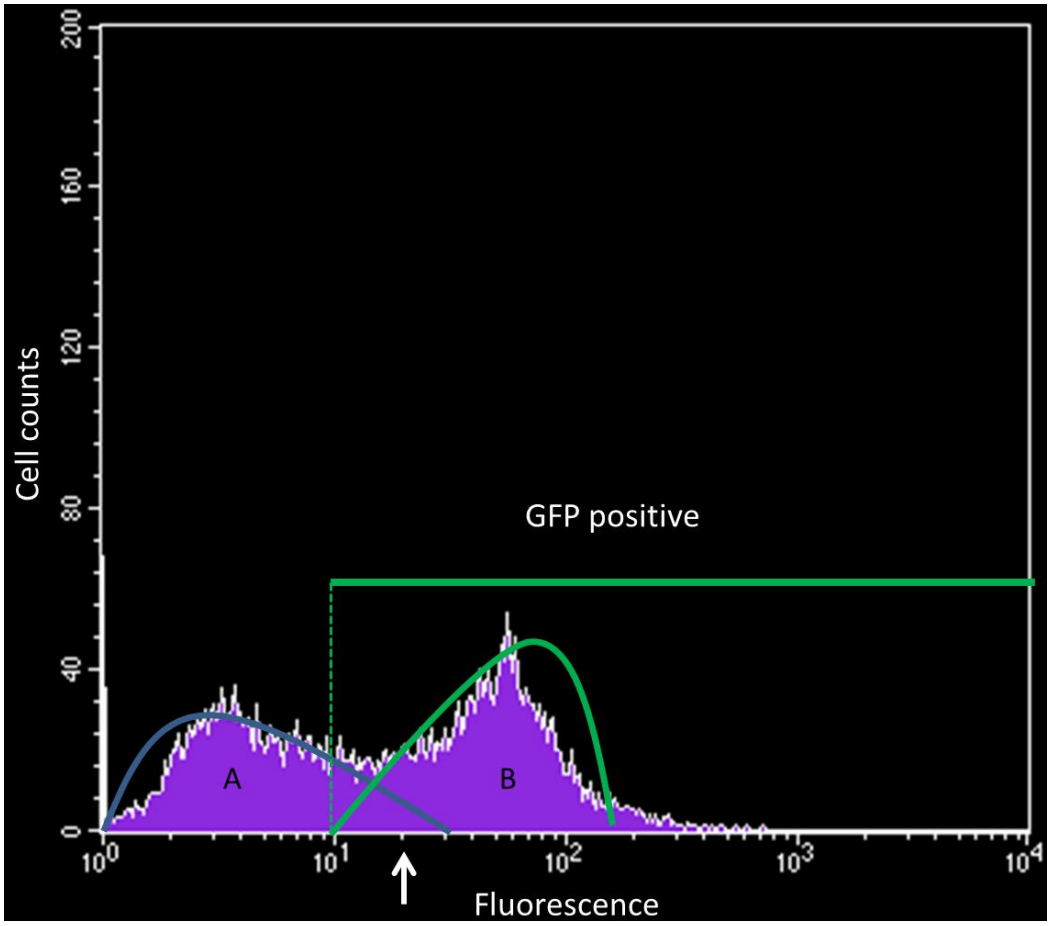


Figure 2.4. Histogram of transfected cells analysed by flow cytometry with cell count plotted against fluorescence intensity

A representative histogram of cells analysed by flow cytometry. Two distinct peaks show A) untransfected cells under the blue curve and B) transfected cells expressing GFP under the green curve. The green line indicates the range of fluorescence intensity for GFP positive cells. The white arrow indicates the mid-point used to assess the percentage of cells which were transfected (Ducrest et al., 2002; Tung et al., 2007).

2.3.6 Functional cAMP assays

Functional assays were carried out using a cAMP ELISA kit (Cell Biolabs). All experiments were carried out in triplicate and independent experiments were repeated three times. Media were removed from the wells and cells were gently washed with PBS. Varying concentrations of α-MSH (Sigma) and ASIP (93-132-amide mouse, Phoenix Peptides) were prepared in serum-free media. Serum-free media (400 µl) containing the relevant concentrations of α-MSH and/or ASIP were added to each well and cells were incubated at 37^{o}C for 45 minutes. The following concentrations of α-MSH and/or ASIP were tested: 0, 0.1, 1, 10, 100, 1000, 10 000, 100 000 nM. Media were then removed and cells were washed carefully with PBS. Ice cold lysis buffer (200 µl) was added to each well and the cells were shaken in plates for 20 minutes on ice. The contents of each well were spun at 13 000 rpm for 10 minutes and cAMP was measured using a colorimetric ELISA kit, following the manufacturer's instructions. Results were normalised against total cellular protein content (see 2.3.8).

2.3.7 ELISA to measure intracellular cAMP

A dilution series of cAMP standard was prepared in lysis buffer to include the following dilutions: 100 µM, 10 µM, 1 µM, 0.1 µM, 0.01µM, 0.001µM, 0.0001 µM, 0 µM.
50 µl of cAMP sample or standard were added to the ELISA plate. Diluted Peroxidase cAMP Tracer Conjugate (25 µl) (1:100 dilution in assay diluent) was added to each test well. Diluted Rabbit Anti-cAMP Polyclonal Antibody (50 µl) (1:500 dilution in assay diluent) was added to each test well. The plate was covered and incubated at room temperature for 2 hours on a rotary shaker. Wells were emptied and each well was washed five times with 250 µl 1 × wash buffer per well, with thorough aspiration between each wash. Substrate solution (100 µl) was added to each well and the plate was incubated at room temperature for 5-20

minutes on a rotary shaker, until colour had developed. Stop solution (100 µl) was added to each well. Absorbance of each microwell was read on a spectrophotometer (Tecan, Sunrise) at 450 nm.

2.3.8 Protein assay

In order to normalise the data from the cAMP assay, a protein assay kit (Biorad) was used to determine total protein levels from cell lysates. Dilutions of protein standard (albumin) were prepared ranging from 0.2 mg/ml to 1.5 mg/ml protein in cell lysis buffer. Standards and samples (25 µl) were pipetted into clean, dry microcentrifuge tubes. RC Reagent I (125 µl) was added to each tube and vortexed. Tubes were incubated for 1 hour at room temperature. RC Reagent II (125 µl) was added to each tube and vortexed. Tubes were centrifuged at 13 000 rpm for 5 minutes. The supernatant was discarded and 127 µl Reagent A´ was added to each tube and vortexed. Tubes were incubated at room temperature for 5 minutes then vortexed. 1 ml of DC Reagent B was added to each tube and vortexed immediately. Tubes were incubated at room temperature for 15 minutes and absorbances were read at 750 nm on a spectrophotometer (Thermo Scientific).

2.4 Computer Analysis

2.4.1 DNA sequence analysis

DNA and amino acid sequences were compared to sequences in the National Center for Biotechnology Information (NCBI) database using the basic local alignment search tool (BLAST) (BLAST, 2014). Alignments were performed using Clustal Omega programmes

(Sievers et al., 2011; EMBL, 2014). Chromatograms were examined by eye to identify heterozygotes using FinchTV (Dymond et al., 2009).

2.4.2 Computer programmes for structural studies

Amino acid sequences were used to predict the structure of the MC1R-wt, MC1RΔ24, ASIP and α-MSH of the grey squirrel. Sequences were uploaded to the web servers Phyre2, I-TASSER and Pepfold in order to generate pdb files giving the three dimensional structures of the molecules under consideration (Zhang, Arakaki and Skolnick, 2005; Kelley and Sternberg, 2009; Thevenet et al., 2012). The pdb files obtained from Phyre2, I-TASSER and Pepfold were analysed further with various computer programmes. The computer programmes Hex 8.0 (Macindoe et al., 2010) and Molsoft (Abagyan, 2012) were used to visually investigate the three dimensional structures. Pdb files were analysed directly using the programme "Socket" (Walshaw and Woolfson, 2001) to investigate intramolecular interactions, and "Fpocket" (Le Guilloux, Schmidtke and Tuffery, 2009) to predict the location of binding pockets on the molecules. Pdb files were uploaded to the web server "Pepsite" (Thevenet et al., 2012) to predict the specific binding between the MC1R receptor and the α-MSH peptide. The protein docking programme *CLC Discovery* (Thomsen and Christensen, 2006) was used for predicting the intermolecular interactions. Docking between the molecules of interest was carried out according to the instruction manual except the default setting of 100 iterations was changed to 200 iterations.

2.5 Statistical Analysis

All data from functional assays in this study were analysed statistically using Student's t-test to evaluate the differences in a) basal levels of cAMP, b) EC_{50} and c) maximal levels of cAMP between each case (see figure 2.5).

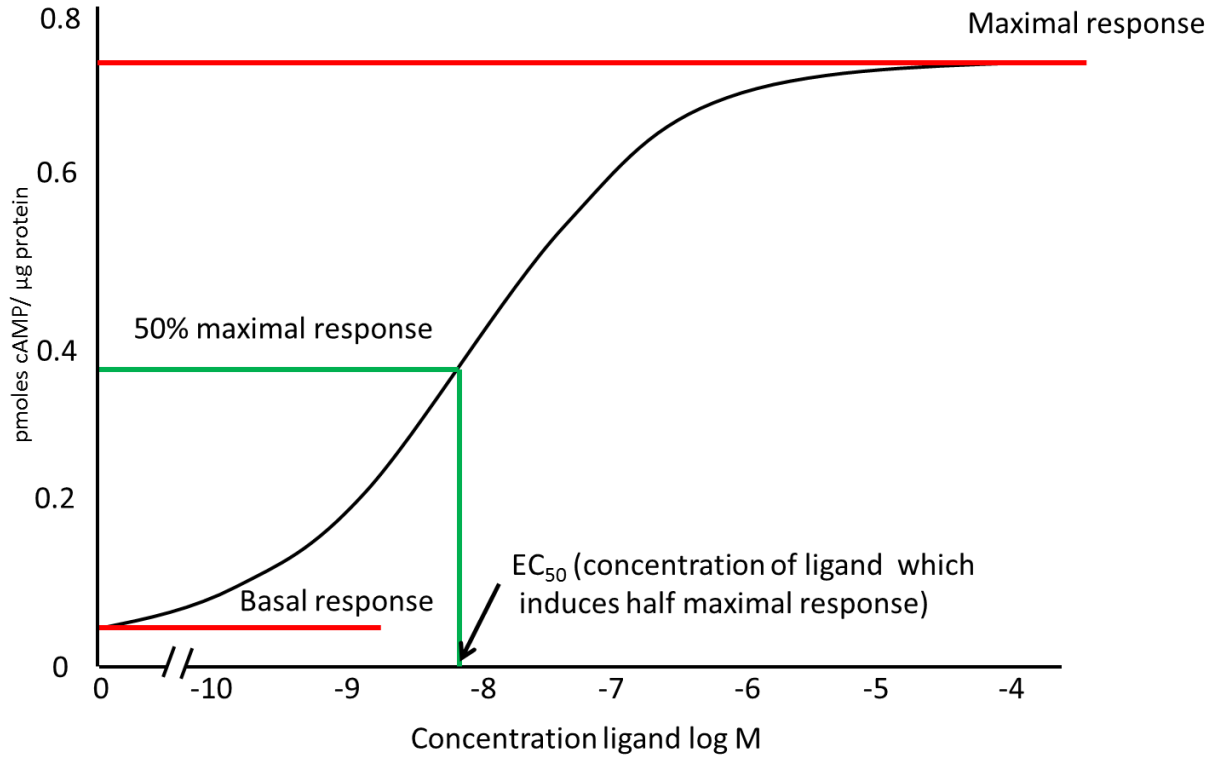


Figure 2.5. Graph to indicate basal and maximal responses and EC_{50} for a typical dose response curve

The null hypothesis in each case was that there is no difference between means of the a) basal response, b) maximal response and c) EC_{50} for the variants under consideration. Three sets of data for each variant were used. The t-test was a one-tailed test and data were related for

assays that were carried out in parallel and unrelated for assays not carried out in parallel. A critical significance level of 5% was used. T-tests were carried out using Microsoft Excel.

Chapter Three

The Melanocortin-1 Receptor of the Grey, Fox and Red Squirrel

3.1 Introduction

Melanism is found in a variety of squirrel species including the grey, fox and red squirrel. All three of these species live in mixed populations where wildtype morphs interbreed freely with melanic morphs of the same species (Gurnell et al., 2004). Variations in the *MC1R* gene are associated with melanism in a wide variety of species including mammals and birds (Robbins et al., 1993; Theron et al., 2001; Majerus and Mundy, 2003). This gene was therefore the first candidate for scrutiny in the investigation of the genetic basis of melanism in the grey, fox and red squirrel. The initial hypothesis was that variations in the *MC1R* gene are associated with melanism in the grey, fox and red squirrel.

The specific methods used in this section included phenotype analysis, genomic DNA extraction, DNA amplification by PCR, sequencing of PCR products and DNA sequencing. Sequences were aligned using BLASTN and Clustal Omega.

3.2 Results

3.2.1 Phenotypes

A number of specimens of wildtype and melanic grey squirrels were available for direct examination of the whole body colouration and of individual hairs. A typical wildtype grey squirrel has an overall grizzled grey appearance with a pale ventrum (figure 3.1 and 3.2). Melanic squirrels are of two distinct phenotypes. The brown-black phenotype has an overall dark brown appearance with an orange ventrum. The jet black phenotype is jet black on both dorsum and ventrum (figure 3.1 and 3.2). A number of different hair types were detected and table 3.1 shows the hair types found in the three phenotypes. All hairs, with the exception of white hairs on the ventrum, were grey at the base.

Figure 3.1. The three phenotypes of the grey squirrel (*Sciurus carolinensis*)

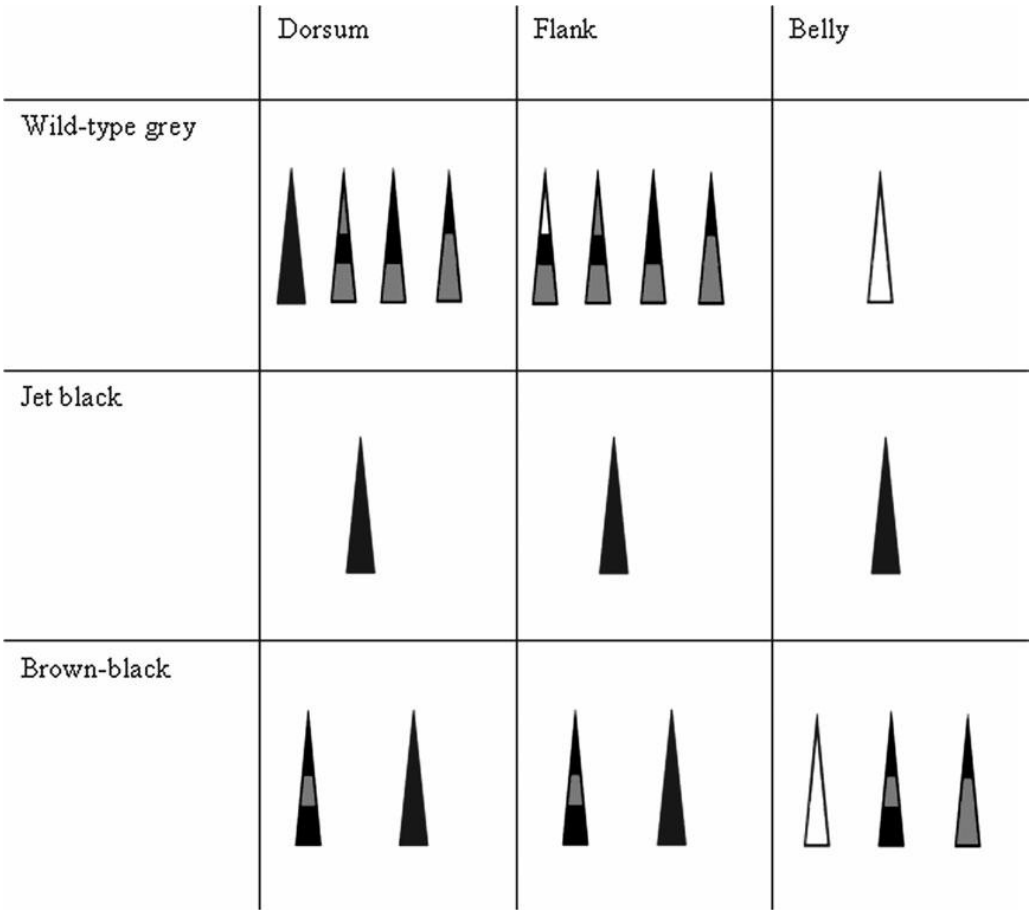
A) Wildtype with inset showing banded hairs, B) jet black and C) brown-black. Photographs by Trimming (2011), Whippey, (2012) and Jaga (2008).

Figure 3.2. Photographs of the three phenotypes of the grey squirrel showing the dorsum and ventrum of each

A) Wildtype dorsum, B) wildtype ventrum, C) brown-black dorsum, D) brown-black ventrum, E) jet black dorsum, F) jet black ventrum. Photographs provided by Dr Nancy Moncrief.

Table 3.1. Diagrammatic representation of hair types from different parts of the body for the wildtype grey, jet black, and brown-black phenotypes of the grey squirrel (*Sciurus carolinensis*)

White indicates little or no pigment (white hair), grey indicates phaeomelanin (orange hair), and black indicates eumelanin (black hair). Table taken from McRobie et al. (2009).



The most common hair types were those depicted in figure 3.3. In the wildtype grey squirrel, the morph comprised a hair with a black tip (average 1.15 mm), an orange band (average 1.95 mm), a black band (average 3.55 mm), an orange band (average 2.5 mm) followed by grey to the root. In the brown-black squirrel the hair comprised of a black tip (average 1.65mm), a shorter orange band (average 0.85 mm) and a black band (average 5 mm) followed by grey to the root. Jet black hairs (type A from table 3.1), were found in all

phenotypes; 3% were type A in the grey squirrel, 10% in the brown-black and 100% in the jet black squirrel.

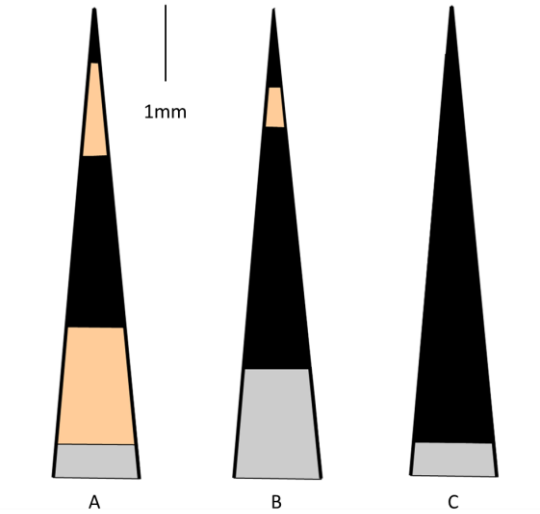


Figure 3.3. Representation of the banded hair of the grey squirrel (*Sciurus carolinensis*)

Hairs from the A) wildtype, B) brown-black and C) jet black phenotypes of the grey squirrel.

No whole specimens were available for examination of the fox squirrel and only small tissue samples were used to extract DNA. Phenotypic information was available from photographs (see figure 3.4) and personal communication with Dr Nancy Moncrief (Virginia Natural History Museum, USA), who kindly donated the samples. The fox squirrel has a number of colour morphs including the wildtype grizzled russet-orange, a dark grey and a completely black phenotype (Moncrief et al., 2010). Only the wildtype and black phenotypes are included in the present study.

Figure 3.4. Two phenotypes of the fox squirrel (*Sciurus niger*)

A) Wildtype fox squirrel B) melanic fox squirrel, C) wildtype dorsum, D) wildtype ventrum, E) melanic dorsum and F) melanic ventrum. Photographs A and B by Loser (2007) and C-F are provided by Dr Nancy Moncrief.

Regarding the red squirrels, only the British samples were available for whole body examination. All Italian samples were provided with a brief description of the phenotype. The red squirrel has a spectrum of colour morphs ranging from russet-red, red-brown, brown, to black. Red squirrels can also be black with a grey dorsum and black with red flanks (Lapini L, personal communication) (See figure 5).

A B

Figure 3.5. Two phenotypes of the red squirrel (*Sciurus vulgaris*)

A) Wildtype and B) melanic red squirrel. Photographs by Trimming (2011) and Whippey (2012).

3.2.2 Genetics

3.2.2.1 Investigation of the *melanocortin-1 receptor* gene of wildtype and melanic grey squirrels

Genomic DNA was used to amplify the *MC1R* gene of both wildtype and melanic grey squirrels. The gene was amplified in two sections, as shown in the primer map of the gene in figure 3.6.

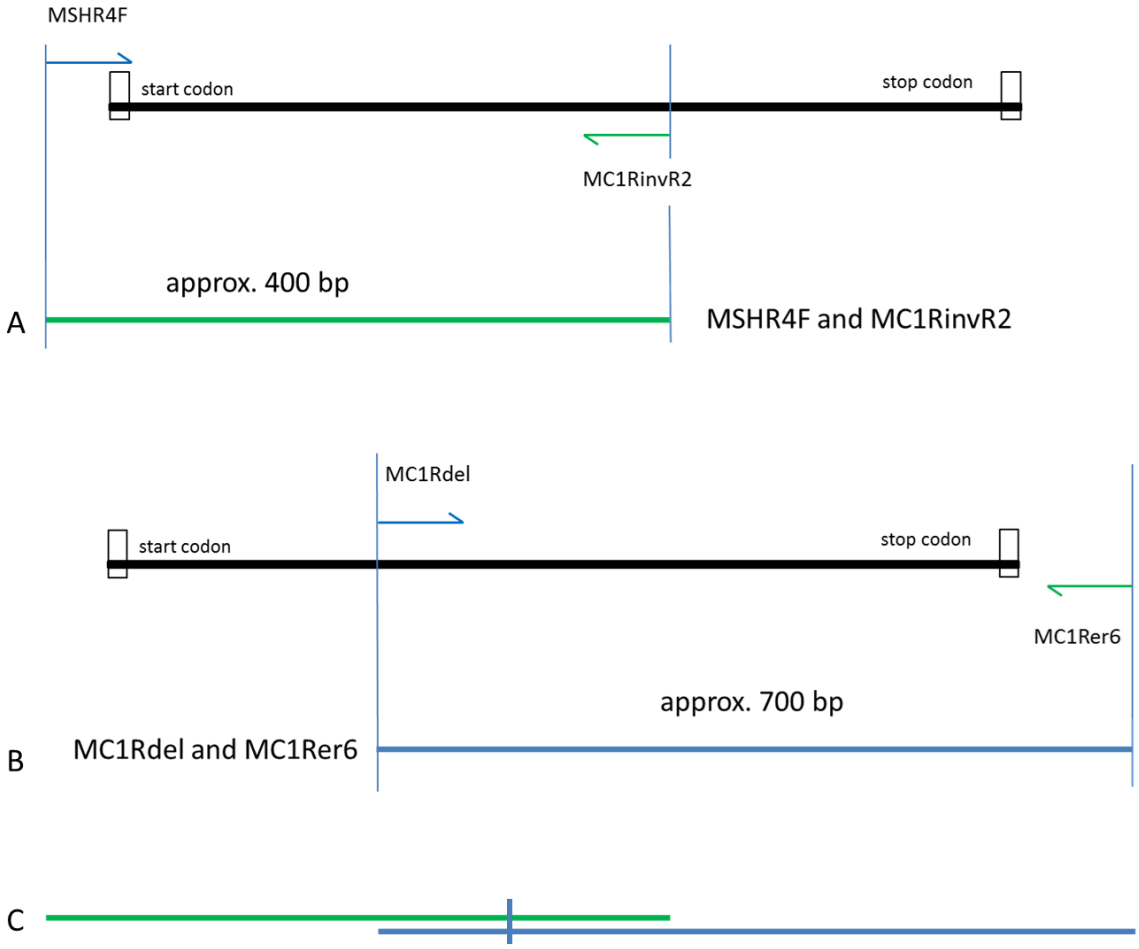


Figure 3.6. Primer map of the *melanocortin-1 receptor* gene

Primer map showing relative positions of the forward and reverse primers used to amplify the MC1R gene. A) The primers MSHR4F and MC1RinvR2 were used to amplify the first section of the gene and B) the primers MC1Rdel and MC1Rer6 were used to amplify the second section of the gene. C) shows a representation of the overlapping sections used to create a contiguous sequence, where the horizontal line shows the end of the first section and beginning of the second section. (Not to scale).

The first section was amplified using the primers MSHR4F and MC1RinvR2 (figure 3.6 A). This PCR product gave a band of approximately 400 bp and figure 3.7 shows a typical agarose gel electrophoresis picture for this reaction.

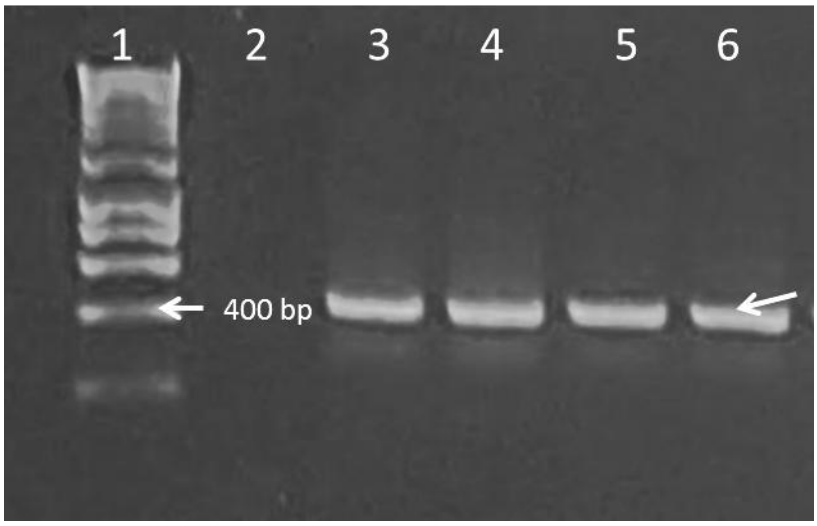


Figure 3.7. PCR amplification of the first section of the *melanocortin-1 receptor* gene from the grey squirrel (*Sciurus carolinensis*)

A representative agarose gel electrophoresis (1.3%) of PCR amplification products visualised with ethidium bromide and illuminated with UV light. Lane 1, Hyperladder 1 (Bioline) DNA ladder and lane 2 is a negative control. Lanes 3, 4, 5 and 6 show 400 bp products of the MC1R of the grey squirrel using the primers MSHR4F and MC1RinvR2. The arrow in lane 6 indicates the band of interest.

This product was purified and sequenced with the MSHR4F forward primer and with the MC1RinvR2. Figure 3.8 shows a typical chromatogram obtained from this PCR reaction. The chromatogram is the reverse-complement of the original sequence as it was sequenced with a reverse primer. The chromatogram shows high quality sequence and includes a single nucleotide polymorphism. The grey bar above the base call gives an indication of the quality of the sequence, with a high bar indicating a base identified with high confidence. The bar gives a quality value (not shown) and this number represents the ability of the software used to identify the base at a particular position. The quality values on the bars shown in this chromatogram are typically between 30 and 80; a value of 30 has an error probability of one

in 1000, and a value of 80, one in 100 000 000 (Geospiza, 2014). A sequence quality value of 30 or above is generally expected by gene sequence databases. The sequence from bp positions 380-402 (figure 3.8), is of poor quality, and is typical of that obtained for the first 50 bp (approximately) from the sequencing primer, as is the case here. The blue line at position 351 indicates the end of the first section of the gene used to create a contiguous sequence of the whole gene. This position corresponds to the vertical blue line on the primer map (figure 3.6C). The red arrow in figure 3.8 highlights a single nucleotide polymorphism (SNP), identified in this particular sequence. SNPs will be considered in more detail 3.2.2.3.

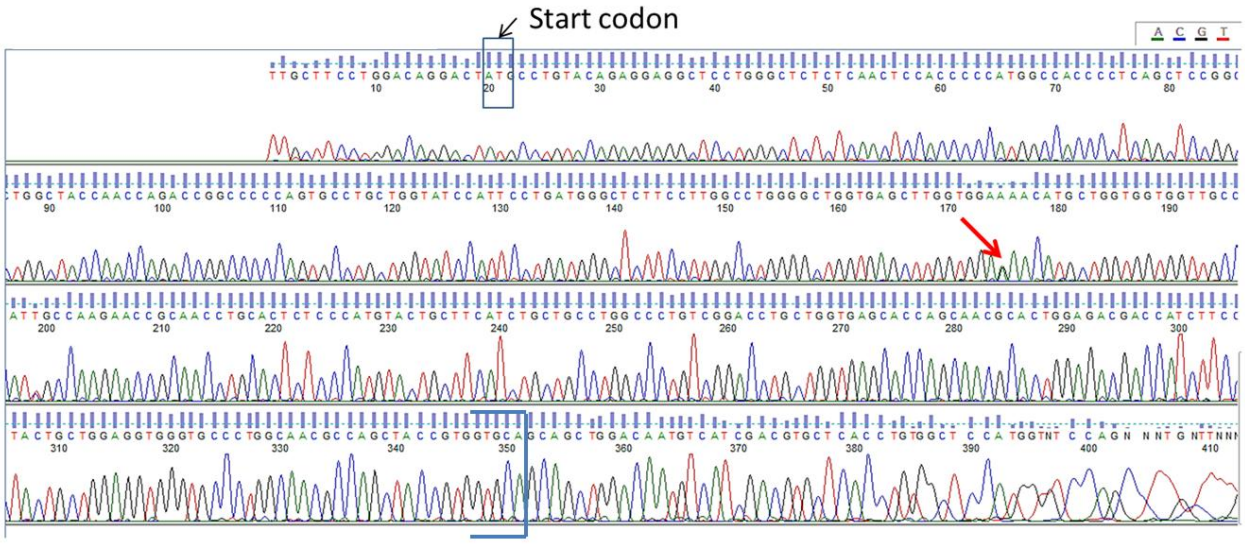


Figure 3.8. Chromatogram of the first section of the *melanocortin-1 receptor*

Chromatogram obtained from sequencing a MSHR4F and MC1RinvR2 PCR product using MC1RinvR2 to sequence. Numbers indicate base pairs counting from the beginning of the sequence. Coloured letters indicate base pairs. The grey bars above the letters are quality bars, where a high bar indicates a base identified with high confidence. The box indicates the start codon. The blue line at position 351 indicates the end of the first section of sequence used for creating the contiguous sequence of the whole gene. The red arrow highlights a single nucleotide polymorphism identified in this sequence.

The second section of the gene was amplified using the primers MC1Rdel and MC1Rer6 (figure 3.6 B). An agarose gel electrophoresis picture for a typical PCR reaction using MC1Rdel and MC1Rer6 is shown in figure 3.9.

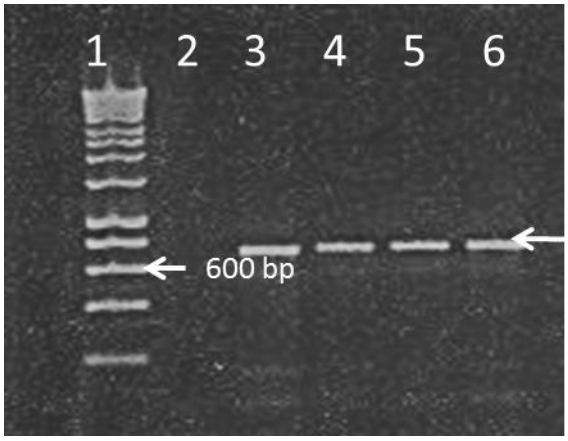


Figure 3.9. PCR amplification of the second section of the *melanocortin-1 receptor* gene from the grey squirrel (*Sciurus carolinensis*)

Representative agarose gel electrophoresis (1.3%) of PCR amplification products visualised with ethidium bromide and illuminated with UV light. Lane 1, Hyperladder 1 (Bioline) DNA ladder and lane 2 is a negative control. Lanes 3, 4, 5 and 6 show 700 bp products of the MC1R of the grey squirrel using the primers MC1Rdel and MC1Rer6. The arrow in lane 6 indicates the band of interest.

Figure 3.10 shows a chromatogram obtained from this PCR product sequenced with the MC1Rdel primer. The blue lines at position 45 show the position of the beginning of the second section of the gene used to create a contiguous sequence of the whole gene and corresponds to the blue line in figure 3.6 C.

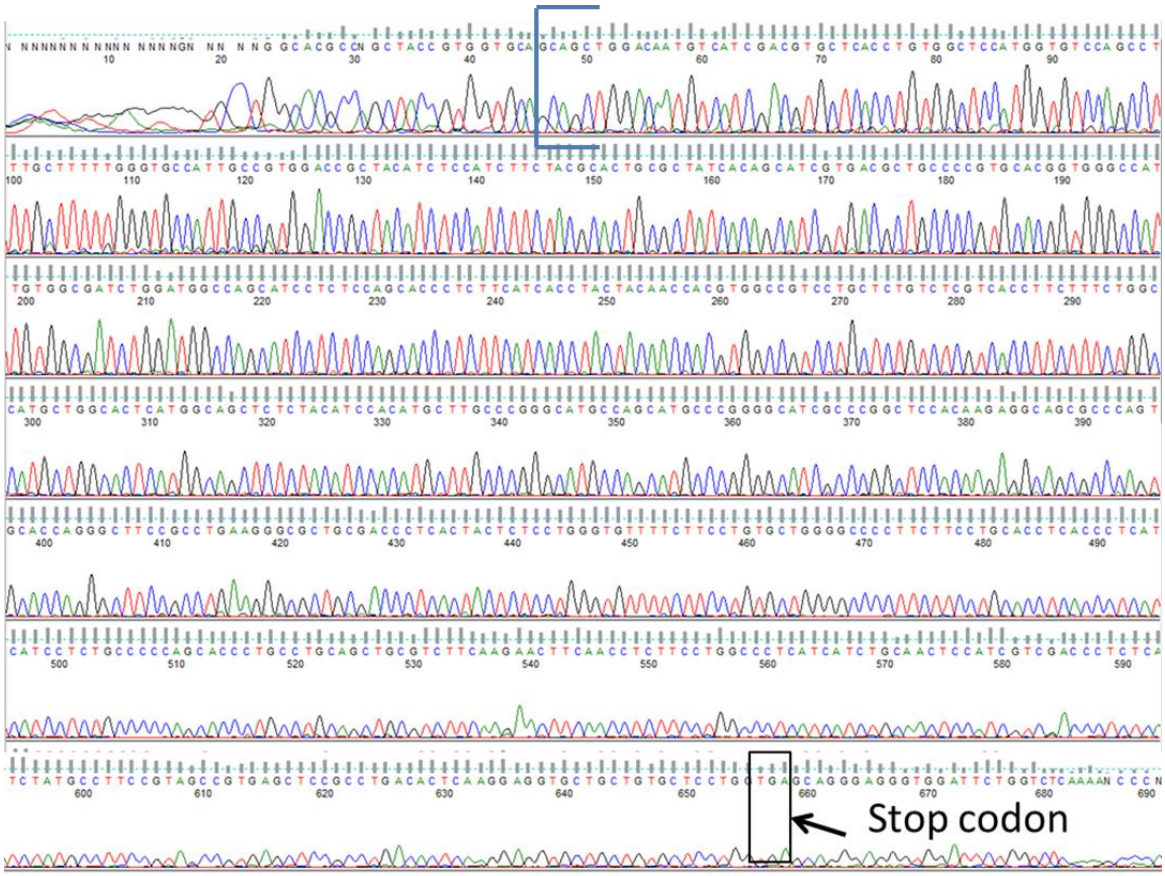


Figure 3.10. Chromatogram of the second section of the *melanocortin-1 receptor*

Chromatogram obtained from sequencing a MC1Rdel and MC1Rer6 PCR product using MC1Rdel to sequence. Coloured letters indicate base pairs. The grey bars above the letters are quality bars, where a high bar indicates a base identified with high confidence. The blue line at position 45 indicates where the second section of the gene starts for creating the contiguous sequence of the whole gene and the box indicates the stop codon.

Sequences from the first and second sections of the gene, shown above, were used to create a contiguous sequence for the *MC1R*, as shown in figure 3.11. Again, the vertical blue line indicates where the first and second sections of the gene sequence join to create the contiguous sequence.

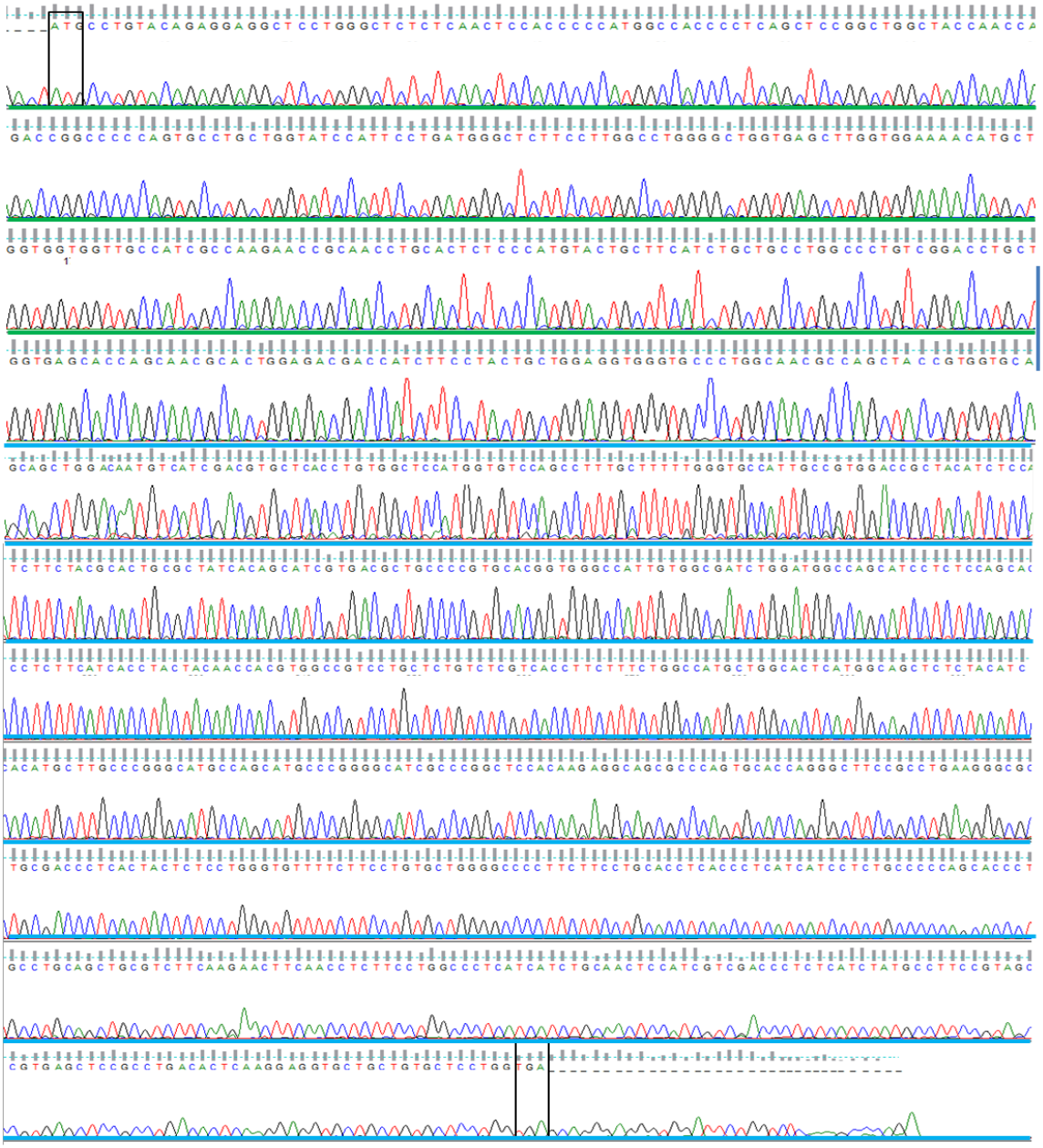


Figure 3.11. Chromatogram of the contiguous sequence of the *melanocortin-1 receptor*

Chromatogram underlined in green is derived from sequencing the MSHR4F and MC1RinvR2 PCR product, using MC1RinvR2 to sequence. Coloured letters indicate base pairs. The grey bars above the letters are quality bars, where a high bar indicates a base identified with high confidence. The vertical blue line at the end of the third row indicates where sequences were joined to create a contiguous sequence of the whole gene. Chromatogram underlined in blue is derived from sequencing the MC1Rdel and MC1Rer6 PCR product, using MC1Rdel to sequence. Boxes indicate the start (ATG) and stop (TGA) codons.

This contiguous sequence was compared to sequences in the Genbank database, using BLASTN (BLASTN, 2014). This search confirmed that the *MC1R* gene had been amplified and alignments with the gene from other species were used to identify the start and stop codons of the *MC1R* of the grey squirrel. The *MC1R* of the grey squirrel has 945 bp and encodes a 314 amino acid protein. The sequence for the grey squirrel has 94% identity with the *MC1R* of the thirteen-lined squirrel (*Spermophilus tridecemlineatus*). Figure 3.12 shows an alignment of the *MC1R* of the wildtype grey squirrel and the thirteen-lined squirrel.

Figure 3.12. Alignment of the *melanocortin-1 receptor* gene of the wildtype grey squirrel (*Sciurus carolinensis*) and the thirteen-lined squirrel (*Spermophilus tridecemlineatus*)

Asterisks under bases indicate conserved base pairs and dots indicate bases with similar properties. Numbers refer to the number of base pairs with 1 being the A of the ATG start codon. The boxes indicate the start and stop codons.

Alignment of the *MC1R* sequences obtained from wildtype and melanic squirrels show a 24 bp in-frame deletion at position 258-281, where A of the start codon is 1 (figure 3.13). Here, *MC1R-wt* denotes the wildtype allele and *MC1RΔ24* denotes the allele with the 24 bp deletion. All wildtype samples were homozygous for the *MC1R-wt* allele, brown-black squirrels were heterozygous, having one *MC1R-wt* and one *MC1RΔ24* allele and jet black squirrels were homozygous for the *MC1RΔ24* allele.

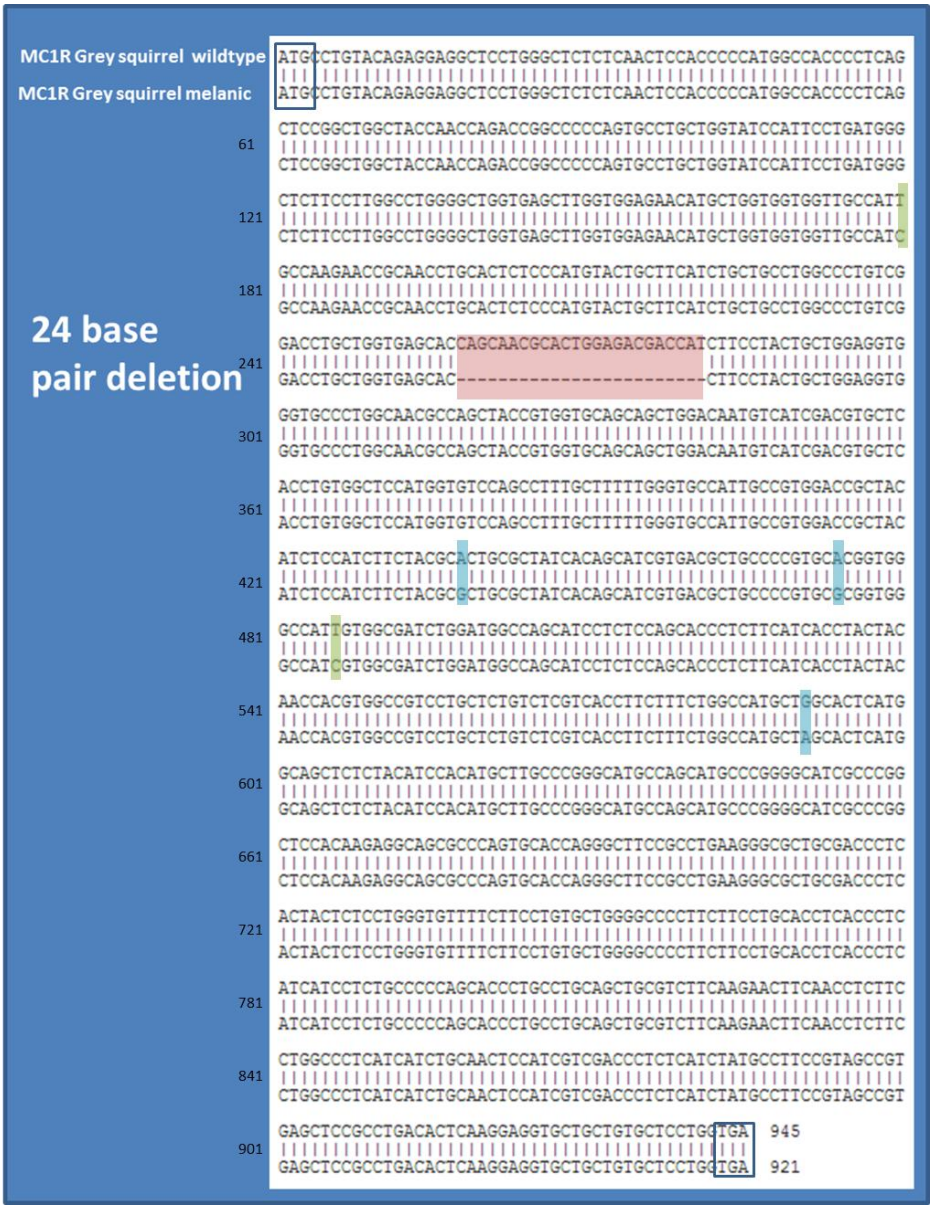


Figure 3.13. Alignment of the *MC1R-wt* and *MC1RΔ24* alleles of the wildtype and melanic grey squirrel (*Sciurus carolinensis*)

DNA sequence of the MC1R-wt (upper sequence) and MC1RΔ24 (lower sequence) of the grey squirrel showing a 24 base pair deletion in the MC1R, associated with melanism, at position 258-281. Lines between base pairs denote consensus between the two sequences. The pale red box highlights the deletion. The green and blue highlighted base pairs are synonymous substitutions. (Blue highlighted base pairs are single nucleotide polymorphisms that are unique to the MC1RΔ24 allele.) Boxes indicate the start and stop codons.

In order to sequence the *MC1R* from heterozygous squirrels, allele specific primers were designed as shown on the primer map in figure 3.14. The primers specific for the *MC1R-wt* allele were designed to anneal to the bases deleted in the *MC1RΔ24* and the primers specific for the *MC1RΔ24* allele were designed to span the deletion. To obtain the sequence of the first section of the gene, PCR products using the primers MSHR4F and MC1RinvR2 were sequenced with either MC1RGR for the *MC1R-wt* allele or MC1RBR for the *MC1RΔ24* allele (figure 3.14 B). To obtain the sequence of the second section of the gene, PCR products using the forward primers MC1Rdel (*MC1R-wt* specific) or MC1RBF (*MC1RΔ24* specific) were used with the reverse primer MC1Rer6.

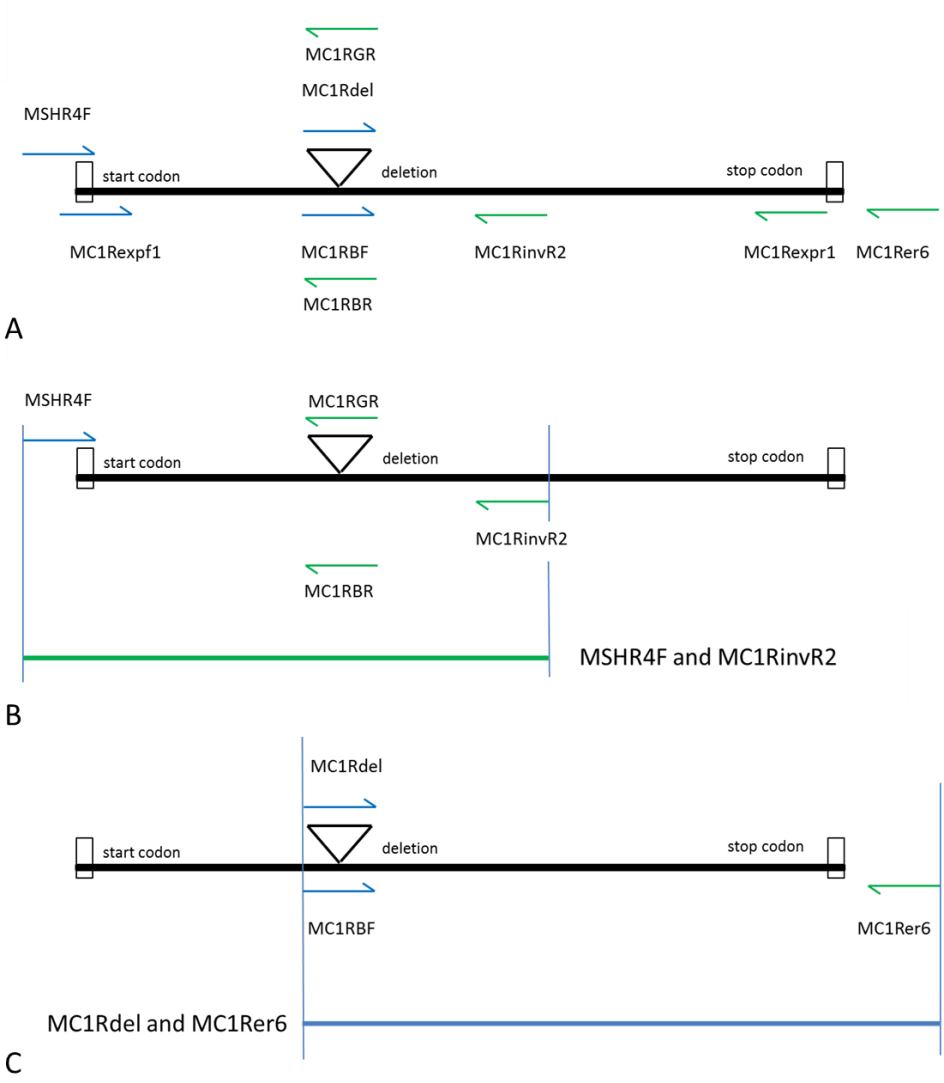


Figure 3.14. Primer map of the *melanocortin-1 receptor* gene

Primer map showing relative positions of the forward and reverse primers used to amplify and sequence the MC1R gene. The triangle represents the 24 base pair deletion in the MC1RΔ24 allele. A) All primers used for amplification and sequencing the MC1R. B) Primers used to amplify the first section of the gene were MSHR4F and MC1RinvR2. The green line represents a PCR product of this reaction. PCR products were sequenced with the allele specific primers MC1RGR or MC1RBR. C) Primers used to amplify the second section of the gene were the allele specific forward primers MC1Rdel or MC1RBF, both with the reverse primer MC1Rer6. The blue line represents a PCR product of this reaction. (Not to scale).

An agarose gel electrophoresis picture for a typical PCR reaction using MC1RBF and MC1Rer6 is shown in figure 3.15. This PCR reaction gave a product of approximately 700 bp.

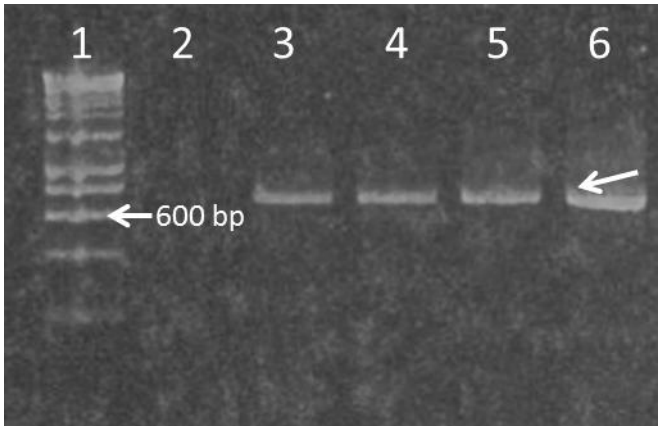


Figure 3.15. PCR amplification of the second section of the *melanocortin-1 receptor* gene from the grey squirrel (*Sciurus carolinensis*) using *MC1RΔ24* allele specific primers

Representative agarose gel electrophoresis (1.3%) of PCR amplification products visualised with ethidium bromide and illuminated with UV light. Lane 1, Hyperladder 1 (Bioline) DNA ladder and lane 2 is a negative control. Lanes 3, 4, 5 and 6 show 700 bp products of the MC1R of the grey squirrel using the MC1RΔ24 allele specific primer MC1RBF and MC1Rer6. The arrow in lane 6 indicates the band of interest.

Figure 3.13 shows an alignment of the *MC1R-wt* and *MC1RΔ24* alleles from wildtype and melanic squirrels respectively. The alignment shows a 24 bp deletion in the *MC1RΔ24,* with bp 258-281 deleted. The *MC1RΔ24* is 921 bp long compared to the *MC1R-wt* which is 945 bp long. Table 3.2 shows the phenotypes and genotypes of wildtype, brown-black and jet black squirrels demonstrating complete association with melanism and the presence of at least one *MC1RΔ24* allele.

Table 3.2. Association between genotype and phenotype in the grey squirrel

Numbers of squirrels found to be of each genotype and corresponding phenotype showing complete association between melanism and the presence of at least one MC1RΔ24 allele.

Phenotype/ Genotype	**Grey**	**Brown black**	**Jet black**
MC1R-wt/ MC1R-wt	51	0	0
MC1R-wt/ MC1RΔ24	0	35	0
MC1RΔ24/ MC1RΔ24	0	0	2

The deletion corresponds to amino acid positions 87-94 in the MC1R protein. The DNA alignment shows a number of SNPs between the two alleles. However, these are synonymous substitutions meaning that the amino acid sequences are unaffected, thus the amino acid sequences of both the MC1R-wt and MC1RΔ24 are identical, except for the eight amino acids deleted. The eight deleted amino acids are serine, asparagine, alanine, leucine, glutamic acid, threonine, threonine and isoleucine (SNALETTI) which give the MC1RΔ24 a 306 amino acid protein. The predicted protein sequences of MC1R-wt and MC1RΔ24 are shown in figure 3.16.

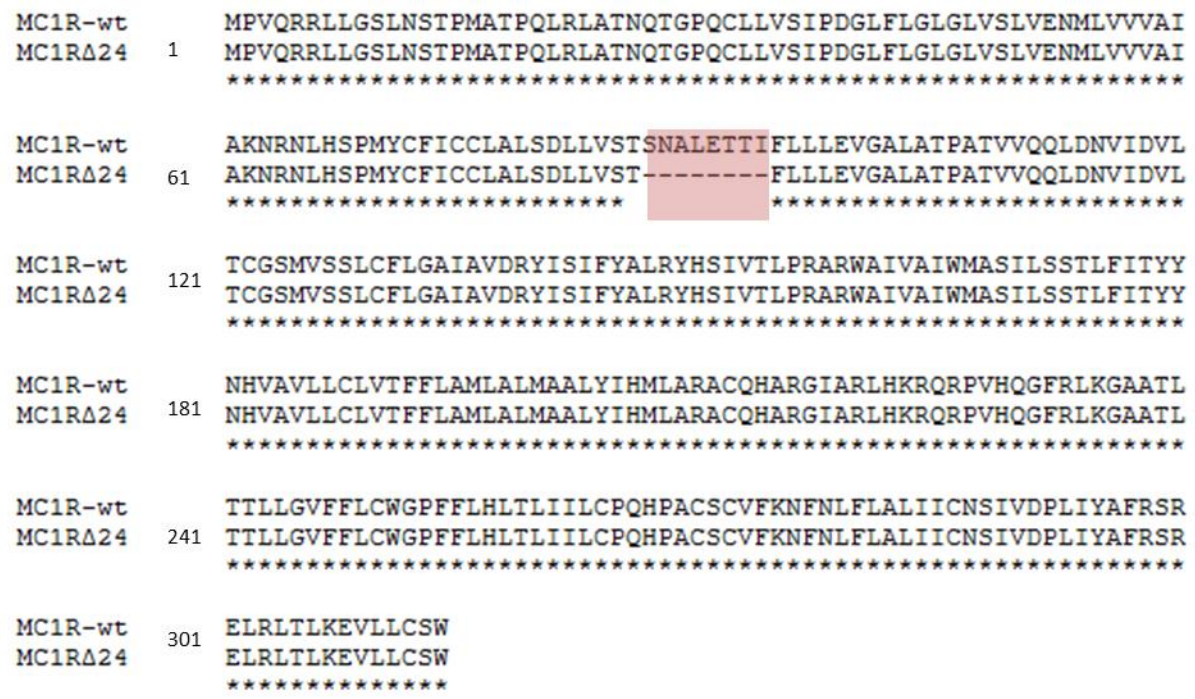


Figure 3.16. Alignment of MC1R-wt and MC1RΔ24 amino acid sequences

The eight amino acids deleted in the MC1RΔ24 are highlighted in pale red at position 87-94. Stars below the sequence denote consensus and dashes indicate deleted amino acids.

3.2.2.2 Origin of the *MC1RΔ24* allele in the grey squirrel

To investigate whether the *MC1RΔ24* allele was a new mutation which had occurred since the introduction of the grey squirrel to the British Isles or whether melanic grey squirrels were introduced from North America, samples of squirrel tissue were obtained and the *MC1R* gene was analysed from all eight melanic North American grey squirrels. The sequences (not shown) show that the melanic grey squirrel of North America and the British Isles has an identical 24 bp deletion in the *MC1R* gene. Of the 44 melanic samples tested, the 42 brown-black samples were heterozygous and the two jet black samples were homozygous for the 24 bp deletion.

3.2.2.3 Alleles of the *melanocortin-1 receptor* in the grey squirrel

Seven SNPs were detected in the *MC1R* gene of the grey squirrel, as shown in table 3.3. All SNPs were synonymous substitutions. In all, five distinct alleles of the *MC1R* were detected in the grey squirrel. These are shown in table 3.3, where *MC1R-wt- G1-4* are wildtype alleles and *MC1RΔ24*-G5 is the melanic allele.

Table 3.3. Alleles of the *melanocortin-1 receptor* of the grey squirrel (*Sciurus carolinensis*)

Single nucleotide polymorphisms (SNP) identified in the MC1R of the grey squirrel compared to the reference sequence which corresponds to the MC1R-wt -G1. The number for each SNP refers to the bp number counting the A of the start codon as 1. Alleles MC1R-wt-G1-4 are alleles of the wildtype grey squirrel and MC1RΔ24-G5 is of the melanic grey squirrel. Accession numbers are shown in brackets. The triangle corresponds to the 24 deleted base pairs.

Base pair/ Allele	156	180	258-281	438	474	486	546	591
MC1R-WT-G1 (EU604831)	G	T	A-C	A	A	C	T	G
MC1R-WT-G2 (KF188574)	A	C	-	-	-	T	C	-
MC1R-WT-G3 (KF188575)	-	C	-	-	-	-	C	-
MC1R-WT-G4 (KF188576)	-	C	-	-	-	-	-	-
MC1RΔ24-G5 (EU60830)	-	C	▲	G	G	-	C	A

Chromatograms of the SNPs are shown in figures 3.17-19. These chromatograms show sections of the *MC1R* sequence with the particular SNP highlighted in both homozygotes and heterozygotes. Figure 3.17 shows the SNPs of the MC1R-wt, figure 3.18 shows another possible SNP of the MC1R-wt, however, this was only detected once and so has not been included in the alleles. Figure 3.19 shows SNPs that were found to be unique to the *MC1RΔ24* allele. These SNPs are also highlighted in blue in the alignment in figure 3.13.

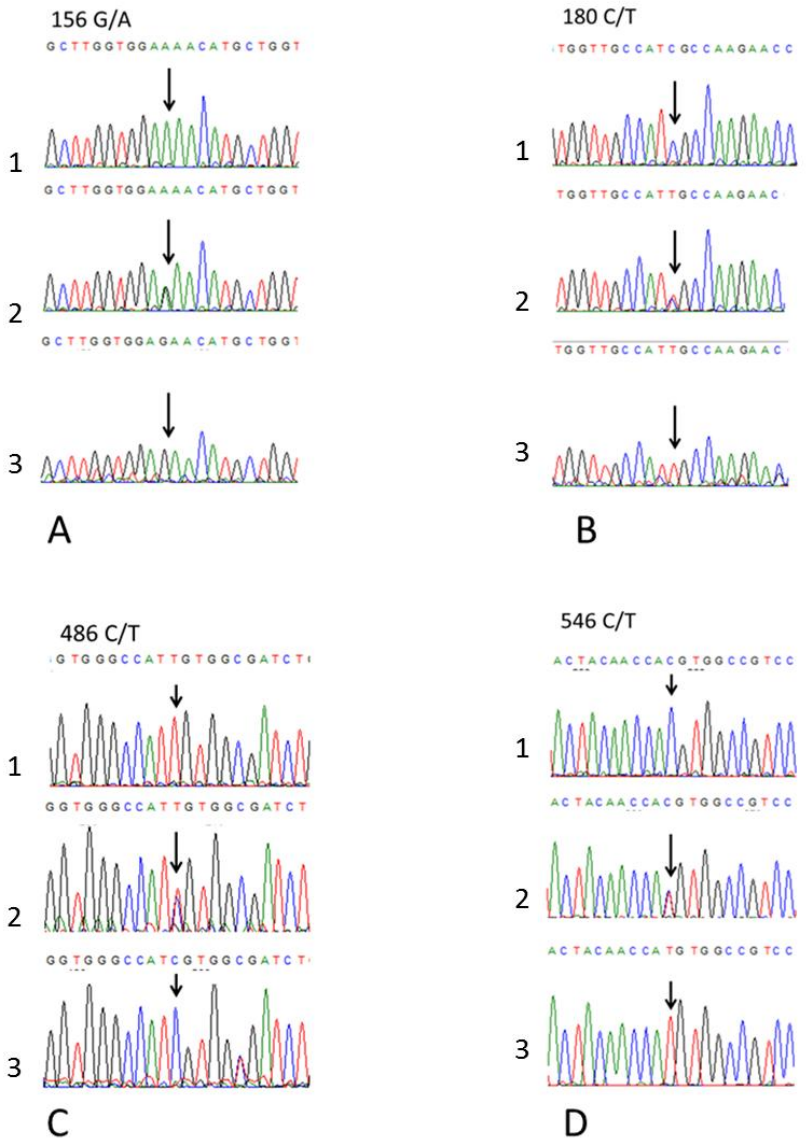


Figure 3.17. Chromatograms showing four different single nucleotide polymorphisms (SNP) of the *MC1R-wt* gene of the grey squirrel (*Sciurus carolinensis*)

Numbers refer to base pair positions where one corresponds to the A of the start codon. The vertical arrow highlights the SNP in each case. A) The 156 G/A SNP with the A homozygote (1), A/G heterozygote (2), the G homozygote (3). B) 180 C/T SNP with the C homozygote (1), C/T heterozygote (2), and the T homozygote (3). C) 486 C/T SNP with the T homozygote (1), C/T heterozygote (2), and the C homozygote (3). D) 546 C/T SNP with the C homozygote (1), C/T heterozygote (2), and the T homozygote (3).

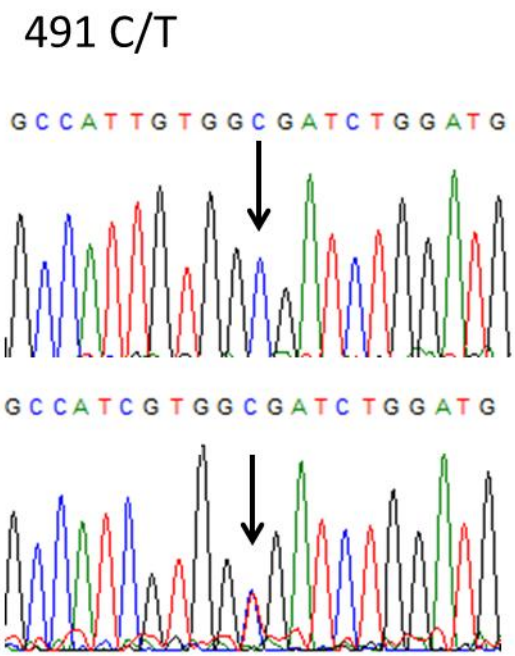


Figure 3.18. Chromatograms showing a possible single nucleotide polymorphisms (SNP) of the *MC1R-wt* allele of the grey squirrel (*Sciurus carolinensis*)

Numbers refer to base pair positions where one corresponds to the A of the start codon. The vertical arrow highlights the SNP in each case. A possible 491 C/T SNP with the C homozygote at the top and C/T heterozygote at the bottom. The T was only detected once and could be an artefact so this SNP has not been included as an allele.

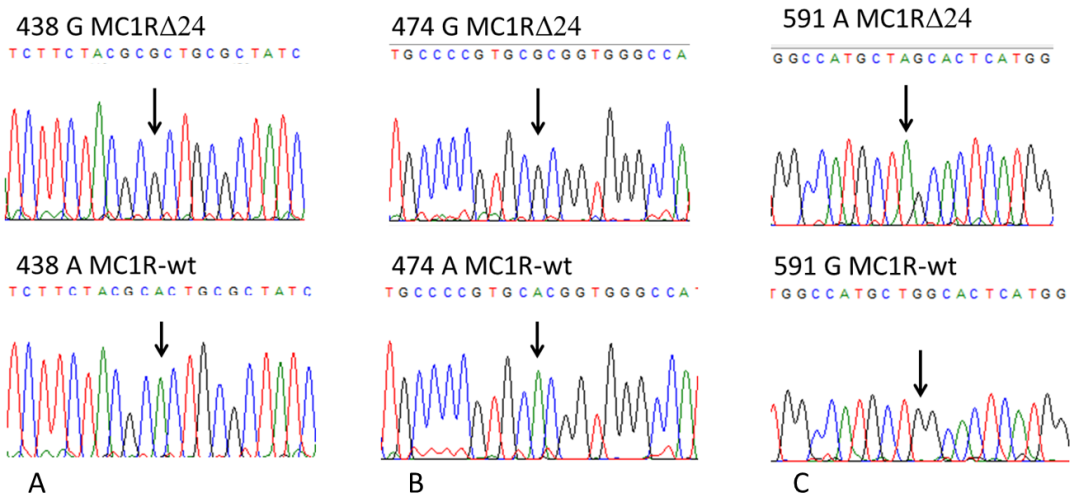


Figure 3.19. Chromatograms showing single nucleotide polymorphisms (SNPs) unique to the *MC1RΔ24* allele of the grey squirrel (*Sciurus carolinensis*)

Numbers refer to base pair positions where one corresponds to the A of the start codon. The vertical arrow highlights the SNP in each case. A) the 438 A/G SNP, B) the 474 A/G SNP and C) the 591 G/A SNP with the MC1R-wt base pair as reference.

An alignment of the four wildtype alleles (MC1R-wt-1-4) is shown in figure 3.20.

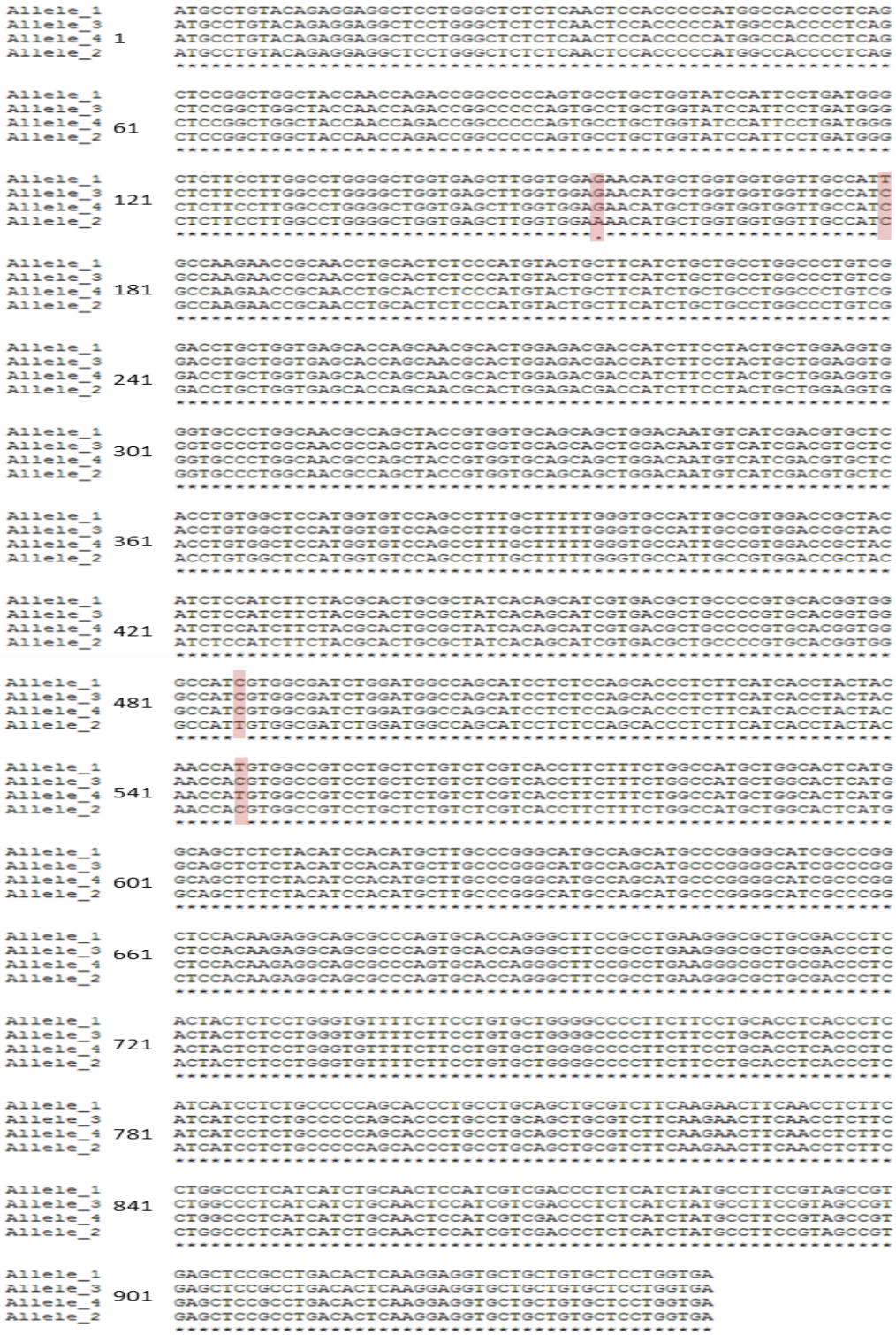


Figure 3.20. Alignment of the four *MC1R-wt* alleles of the grey squirrel (*Sciurus carolinensis*)

Asterisks under base pairs denote consensus between sequences. The base pairs highlighted in pale red are single nucleotide polymorphisms. Numbers refer to the number of base pairs with 1 being the A of the ATG start codon.

All five alleles were detected in the British and North American samples. Interestingly, three of the SNPs identified are unique to the *MC1RΔ24* allele and even though the melanic squirrels tested were from diverse locations, they all contained these SNPs. Allele counts found in different locations in Britain and North America are shown in table 3.4 and figure 3.21.

Table 3.4. Total number of squirrel samples, allele counts, frequencies and phenotypes of each location from Britain and North America

The numbers under phenotype refer to the numbers of squirrels. The numbers under alleles refer to allele counts (top) and frequencies (bottom) and phenotypes. wt= wildtype, bb= brown black, jb= jet black, Northants = Northamptonshire, W. Yorks = West Yorkshire, Cambs = Cambridgeshire, BC = British Columbia, MA = Massachusetts, VA = Virginia. (It should be noted that frequencies comparing the MC1R-wt with MC1RΔ24 alleles are not representative of the populations as melanic squirrels were selected preferentially over wildtype squirrels for this study.)

Population	Phenotype			Alleles										Total	
	Grey	Brown black	Jet black	G1 wt	G1 bb	G2 wt	G2 bb	G3 wt	G3 bb	G4 wt	G4 bb	G5 bb	G5 jb	MC1R-wt	MC1RΔ24
Cumbria UK	4	-	-	8	-	-	-	-	-	-	-	-	-	8	-
(frequency)				1	-	-	-	-	-	-	-	-	-		
Northants UK	7	-	-	3	-	5	-	4	-	2	-	-	-	14	-
(frequency)				0.2	-	0.4	-	0.3	-	0.1	-	-	-		
W. Yorks UK	6	-	-	4	-	6	-	2	-	-	-	-	-	12	-
(frequency)				0.3	-	0.5	-	0.2	-	-	-	-	-		
Cambs UK	28	35	1	15	11	16	12	12	8	13	4	35	2	91	37
(frequency)				0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0.3			
BC USA	3	5	-	3	3	2	2	1	-	-	-	5	-	11	5
(frequency)				0.2	0.2	0.1	0.1	0.1	-	-	-	0.3	-		
MA USA	-	-	1	-	-	-	-	-	-	-	-	-	2	-	2
(frequency)				-	-	-	-	-	-	-	-	-	1		
VA USA	3	2	-	1	1	3	1	-	-	2	-	2	-	8	2
(frequency)				0.1	0.1	0.3	0.1	-	-	0.2	-	0.2	-		

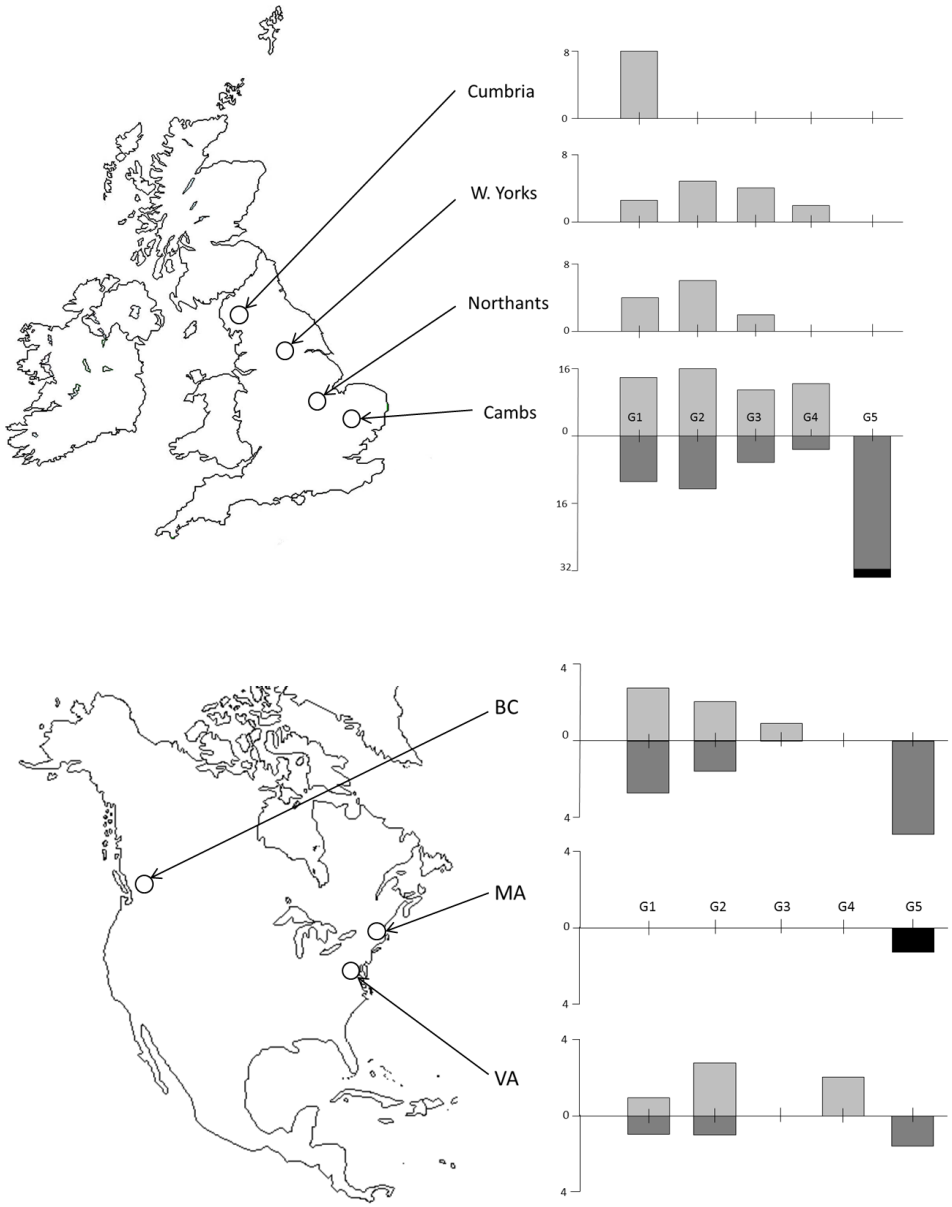


Figure 3.21. Melanocortin-1 receptor allele counts of the grey squirrel (*Sciurus carolinensis*) from the British Isles and North America

G1-G4 are MC1R-wt alleles and G5 is the MC1RΔ24 allele. Allele counts are plotted upwards for wildtype specimens and plotted downwards for melanic specimens. Bars are shaded light grey, dark grey and black for alleles originating from wildtype, brown-black and jet black phenotypes respectively.

3.2.2.4 Investigation of the *melanocortin-1 receptor* gene of wildtype and melanic fox and red squirrels

Given the clear association between the 24 bp deletion and melanism in the grey squirrel, the study was extended to investigate variations in the *MC1R* of the fox squirrel and red squirrel in wildtype and melanic squirrels. No association between variation in the *MC1R* and melanism was found for either the fox or red squirrel. The *MC1R* of the fox squirrel is 945 bp long resulting in a receptor of 314 amino acids. Four SNPs were found and two alleles detected: *MC1R-F1* and *MC1R-F2* (accession numbers KF052119 and KF052120 respectively). A total of nine fox squirrels samples were tested (four wildtype and five melanic). Two of these nine fox squirrel samples were found to be heterozygous for both alleles; all other samples were homozygous. Of the four SNPs detected, one was synonymous and three were non-synonymous substitutions. (Non-synonymous substitutions lead to changes in amino acid sequence.) The substitutions were as follows with the *MC1R-wt* of the grey squirrel as reference: allele 1: P15S and M167I; allele 2: P30S. All amino acid changes are shown in figure 3.25 and all SNPs are shown in table 3.8. No observable differences in phenotype were identified between these alleles. The DNA sequence of the fox squirrel was 99.7% similar to that of the grey squirrel, as shown in the alignment in figure 3.22. Table 3.5 shows a comparison of the SNPs found in *MC1R* of the fox and grey squirrels.

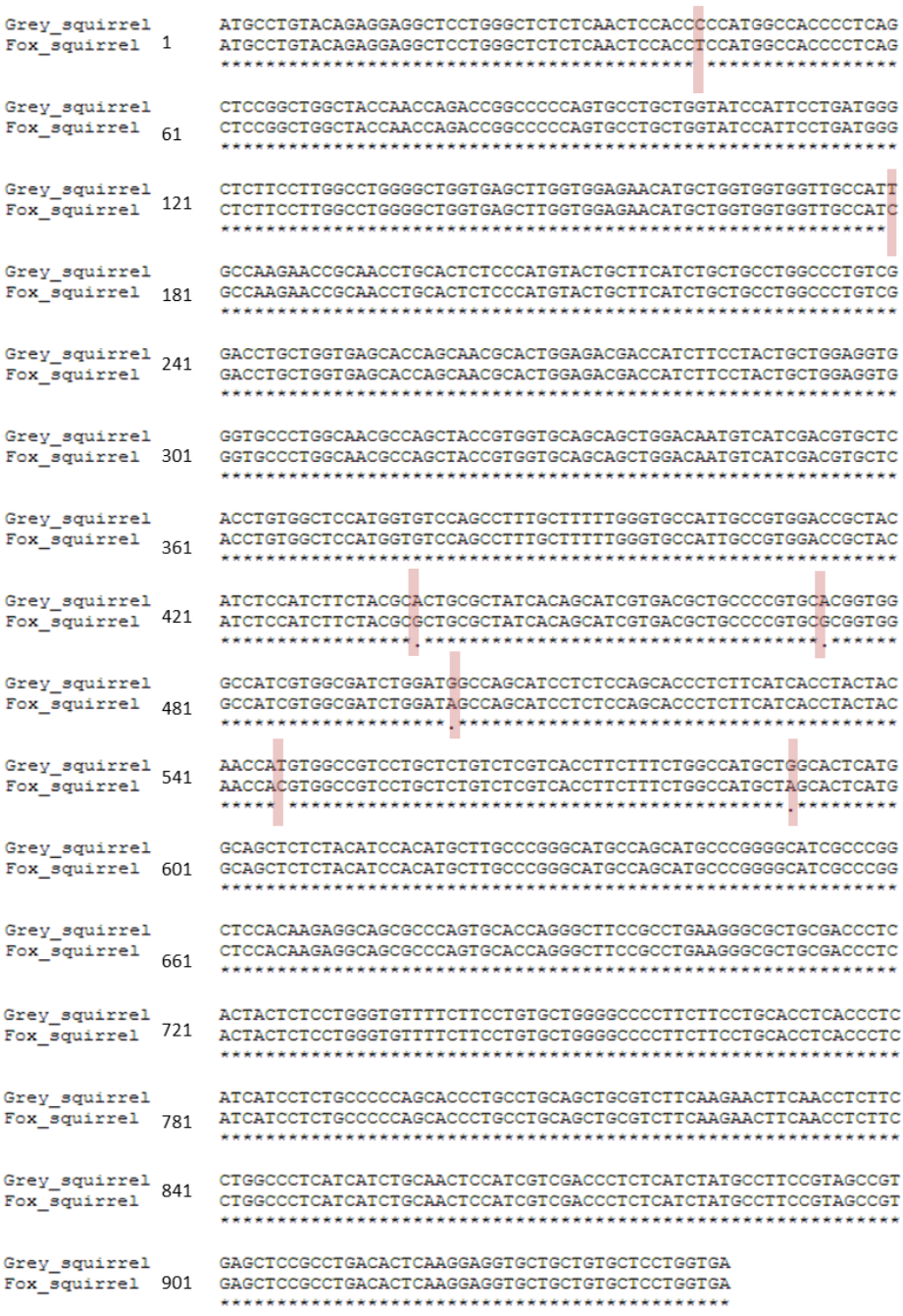


Figure 3.22. Alignment of the *melanocortin-1 receptor* gene of the grey squirrel (*Sciurus carolinensis*) and fox squirrel (*Sciurus niger*)

The grey squirrel sequence is MC1R-wt and the fox squirrel sequence is allele 1. Asterisks under base pairs denote consensus between sequences and dots denote bases with similar properties. The base pairs highlighted in pale red are single nucleotide polymorphisms. Numbers refer to the number of base pairs with 1 being the A of the ATG start codon.

Table 3.5. Comparison of single nucleotide polymorphisms of the melanocortin-1 receptor from the grey (*Sciurus carolinensis*) and fox (*Sciurus niger*) squirrel

The reference sequence corresponds to the MC1RΔ24 allele and the number for each SNP refers to the base pair number counting the A of the start codon as 1 (MC1R-wt numbering). Alleles MC1R-wt G1-G5 are alleles of the grey squirrel and alleles MC1R-F1 and F2 are of the fox squirrel. Asterisks refer to non-synonymous sites. Amino acid change and number following MC1R-wt-G1: 43- S15P; 88- P30S; 501- I167M. The triangle corresponds to deleted base pairs. Total refers to the total number of base pair differences for each allele compared to the reference sequence.

Base pair	43*	88*	105	156	180	259-282	438	474	486	501*	546	591	Total
MC1RΔ24 G5	C	C	A	G	C	▲	G	G	C	G	C	A	-
MC1R-F1	T	-	-	-	-	A-C	-	-	-	A	-	-	2
MC1R-F2	-	T	G	-	-	A-C	-	-	-	-	-	-	2
MC1R-WT G3	-	-	-	-	-	A-C	A	A	-	-	-	G	3
MC1R-WT G4	-	-	-	-	-	A-C	A	A	-	-	T	G	4
MC1R-WT G2	-	-	-	A	-	A-C	A	A	T	-	-	G	5
MC1R-WT G1	-	-	-	-	T	A-C	A	A	-	-	T	G	5

Two alleles of the *MC1R* of the red squirrel were also identified. The first allele (MC1R-R1) (accession number KF188571) has 942 bp giving a 314 amino acid receptor with the following substitutions compared with the wild-type grey squirrel MC1R: S10C, L21F, T105M, T108A, A158V, R233C, and A238V. The second allele (MC1R-R2) (accession

number KF188572) has 939 bp giving a 313 amino acid receptor with the same substitutions as allele 1, compared with the grey squirrel, but also a single amino acid deletion, Y180del. Positions of these amino acids are presented in table 3.6 (and also shown on the schematic diagram of the MC1R in Figure 3.25 in the discussion). Out of a total of 39 red squirrel samples tested, all 33 Italian red squirrel samples (both melanic and wild type) were homozygous for allele 1, while five of the six British samples of the red squirrel (all wild type) were homozygous for the allele 2 and the other was heterozygous.

Table 3.6. Variations in amino acid sequences of the melanocortin-1 receptor between grey, fox and red squirrels with MC1R-wt as the reference sequence

Dashes denote consensus with the reference sequence MC1R-wt of the grey squirrel. Black rectangles represent deleted amino acids. MC1R-F1 and MC1R-F2 are alleles one and two of the fox squirrel respectively and MC1R-R1 and MC1R-R2 are alleles one and two of the red squirrel respectively.

		Amino acids											
Species	Allele	10	15	21	30	87-94	105	108	158	167	180	233	238
Grey squirrel	MC1R-wt	**S**	**P**	**L**	**P**	**SNALETTI**	**T**	**T**	**A**	**M**	**Y**	**R**	**A**
	MC1RΔ24	-	-	-	-	■	-	-	-	-	-	-	-
Fox squirrel	MC1R-F1	-	S	-	-	-	-	-	-	I	-	-	-
	MC1R-F2	-	-	-	S	-	-	-	-	-	-	-	-
Red squirrel	MC1R-R1	C	-	F	-	-	M	A	V	-	-	C	V
	MC1R-R2	C	-	F	-	-	M	A	V	-	■	C	V

3.3 Discussion

3.3.1 Genetic basis of melanism in the grey squirrel

Three distinct colour morphs of the grey squirrel were distinguished; the wildtype grey, brown-black and jet black. The brown-black and jet black are both considered to be melanic, having an overall dark appearance, as shown in figure 3.2C. Melanism in the grey squirrel was found to be perfectly associated with a 24 bp in-frame deletion in the *MC1R* gene, where the wildtype phenotype is homozygous for the *MC1R-wt* allele, the jet black is homozygous for the *MC1RΔ24* allele and the brown-black is heterozygous, having one *MC1R-wt* and one *MC1RΔ24* allele (table 3.2). The jet black morph has solid eumelanic hairs, showing no banding patterns whereas the brown-black morph retains banding patterns of the wildtype but with an overall darkening of the fur. This heterozygote has an intermediate phenotype and therefore the *MC1RΔ24* allele is considered to be incompletely dominant to the *MC1R-wt* allele. The 24 bp deletion results in an eight amino acid deletion which is predicted to fall at the extracellular side of the second transmembrane domain, as shown in figure 3.23. This deletion is close to many other mutations associated with melanism, as shown in figures 1.17 in the introduction and table 3.6. Indeed, the jaguar (Eizirik et al., 2003), jaguarundi (Eizirik et al., 2003), golden-headed lion tamarin (Mundy and Kelly, 2003) and Eleonora's falcon (Gangoso et al., 2011) all have deletions in the same region of the receptor, as shown in figure 1.18 in the introduction. The 24 base-pair deletion in the jet black squirrel corresponds to eight amino acids: serine (polar uncharged), asparagine (polar uncharged), alanine (hydrophobic), leucine (hydrophobic), glutamic acid (negatively charged), threonine (polar uncharged), threonine (polar uncharged) and isoleucine (hydrophobic).

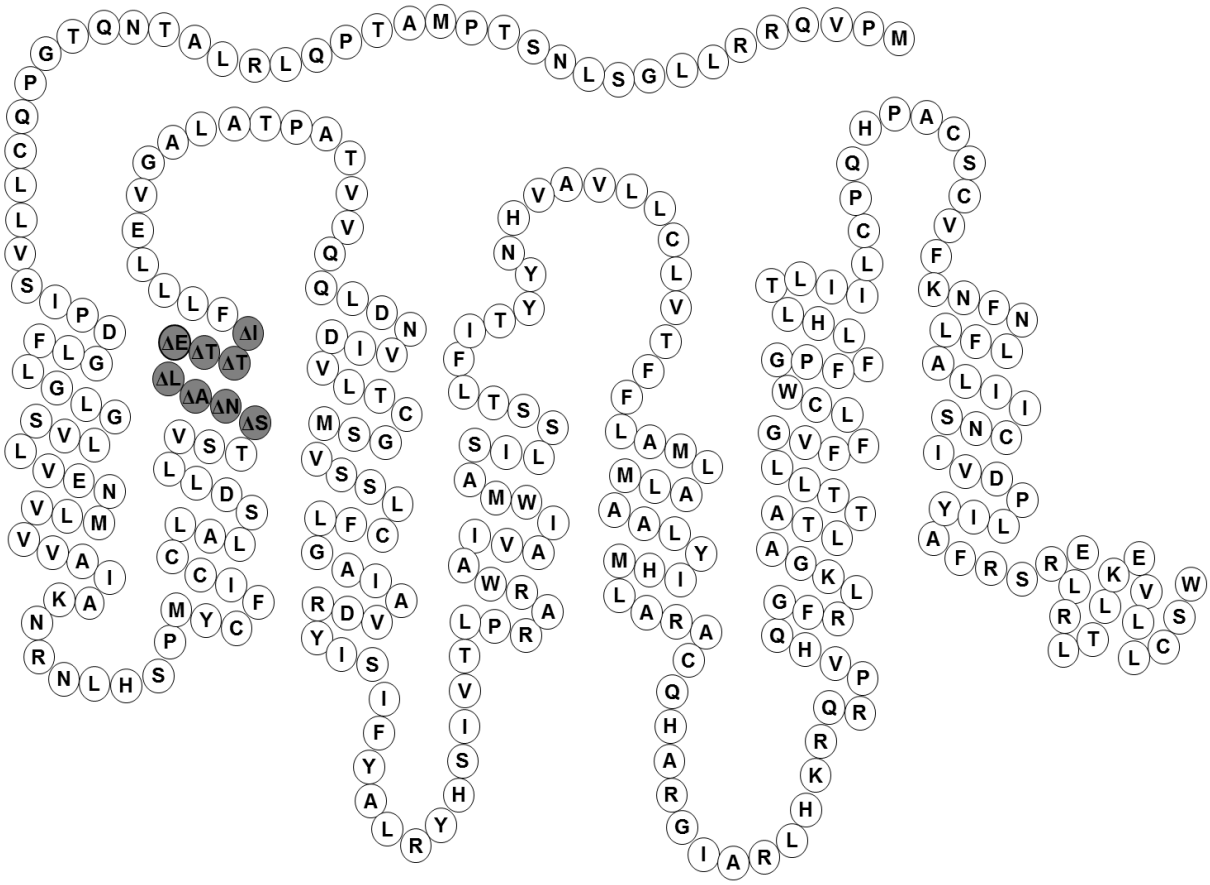


Figure 3.23. Schematic representation of the predicted structure of the melanocortin-1 receptor of the grey squirrel (*Sciurus carolinensis*)

Circles represent amino acids, and dark circles indicate the eight amino acids deleted in the melanic variant.

Table 3.7. Amino acid alignment of melanocortin-1 receptor wildtype and melanic variants of various species

The alignment is taken from amino acids 65-120 (human numbering, sequence not shown). Wildtype and melanic amino acid sequences of the squirrel, pig, mouse, cattle, jaguar, jaguarundi, rabbit, bananaquit, chicken, and Japanese quail are shown. The wildtype (wt) of each species is shown in contrast to the melanic variant underneath. Dashes indicate agreement with the consensus sequence used (squirrel MC1R-wt). Bold letters with asterisks indicate substitutions associated with melanism, and triangles indicate deletions. Transmembrane domains are indicated by boxes. The second transmembrane domain begins at amino acid position 71 and ends at position 99. Source of sequences are as follows: pig (Kijas et al., 1998), mouse (Robbins et al., 1993), (Kijas et al., 1998), jaguar and jaguarundi (Eizirik et al., 2003), rabbit (Fontanesi et al., 2010), bananaquit (Theron et al., 2001), (Ling et al., 2003) and Japanese quail (Nadeau, Minvielle and Mundy, 2006). Table taken from McRobie et al. (2009).

								70										80
Squirrel- E^{+}				K	N	R	N	L	H	S	P	M	Y	C	F	I	C	C	L
Squirrel- E^{a}				-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Pig- MC1R*1				-	-	-	-	-	-	-	-	-	-	Y	-	V	-	-	-
Pig- MC1R*2				-	-	-	-	-	-	-	-	-	-	Y	-	V	-	-	-
Mouse- E^{+}				-	-	-	-	-	-	-	-	-	-	Y	-	-	-	-	-
Mouse- E^{wo}				-	-	-	-	-	-	-	-	-	-	Y	-	-	-	-	-
Mouse- E				-	-	-	-	-	-	-	-	-	-	Y	-	-	-	-	-
Cattle- E^{+}				-	-	-	-	-	-	-	-	-	-	Y	-	-	-	-	-
Cattle- ED				-	-	-	-	-	-	-	-	-	-	Y	-	-	-	-	-
Jaguar- wt				-	-	-	-	-	-	-	-	-	-	Y	-	-	-	-	-
Jaguar- mel				-	-	-	-	-	-	-	-	-	-	Y	-	-	-	-	-
Jaguarundis- red				-	-	-	-	-	-	-	-	-	-	Y	-	-	-	-	-
Jaguarundis- dark				-	-	-	-	-	-	-	-	-	-	Y	-	-	-	-	-
Rabbit- wt				-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Rabbit- E^{D} (black)				-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Bananaquit- Y (yellow)				-	-	-	-	-	-	-	-	T	-	Y	-	-	-	-	-
Bananaquit- M (black)				-	-	-	-	-	-	-	-	T	-	Y	-	-	-	-	-
Chicken- wt				-	-	-	-	-	-	-	-	-	-	Y	-	-	-	-	-
Chicken- black				-	-	-	-	-	-	-	-	T	-	Y	-	-	-	-	-
Japanese quail- wt				-	-	-	-	-	-	-	-	-	-	Y	-	-	-	-	-
Japanese quail- black				-	-	-	-	-	-	-	-	-	-	Y	-	-	-	-	-

									90									100
Squirrel- E^{+}	A	L	S	D	L	L	V	S	T	S	N	A	L	E	T	T	I	F	**L**	L
Squirrel- E^{a}	-	-	-	-	-	-	-	-	-	▲	▲	▲	▲	▲	▲	▲	▲	-	-	-
Pig- MC1R*1	-	V	-	-	-	-	-	-	V	-	-	V	-	-	-	A	V	L	-	-
Pig- MC1R*2	-	V	-	-	-	-	-	-	V	-	-	**M***	-	-	-	A	V	L	**P***	-
Mouse- E^{+}	-	-	-	-	-	M	-	-	V	-	I	V	-	-	-	-	-	I	-	-
Mouse- E^{wo}	-	-	-	-	-	M	-	-	V	-	I	V	-	-	-	-	-	I	-	**P***
Mouse- E	-	-	-	-	-	M	-	-	V	-	I	V	-	**K***	-	-	-	I	-	-
Cattle- E^{+}	-	V	-	-		-	-	-	V	-	-	V	-	-	-	A	V	M	-	-
Cattle- ED	-	V	-	-	-	-	-	-	V	-	-	V	-	-	-	A	V	M	**P***	-
Jaguar- wt	-	V	-	-	-	-	-	-	V	-	S	V	-	-	-	A	V	M	-	-
Jaguar- mel	-	V	-	-	-	-	-	-	V	-	S	V	-	-	-	A	V	M	-	-
Jaguarundis- red	-	V	-	-	-	-	-	-	V	-	S	V	-	-	-	A	V	M	-	-
Jaguarundis- dark	-	V	-	-	-	-	-	-	V	-	S	V	-	-	▲	▲	▲	▲	▲	▲
Rabbit- wt	-	-	-	-	-	-	-	-	V	-	S	V	-	-	-	A	V	L	-	-
Rabbit- E^{D}	-	-	-	-	-	-	-	-	V	-	S	V	-	▲	▲	A	V	L	-	-
Bananaquit- Y (yellow)	-	V	-	-	M	-	-	-	I	-	-	L	A	-	M	L	F	M	-	-
Bananaquit- M (black)	-	V	-	-	M	-	-	-	I	-	-	L	A	**K***	M	L	F	M	-	-
Chicken- wt	-	V	-	-	M	-	-	-	V	-	-	L	A	-	-	L	F	M	-	-
Chicken- black	-	V	-	-	M	-	-	-	V	-	-	L	A	**K***	-	L	F	M	-	-
Japanese quail-wt	-	V	-	-	M	-	-	-	V	-	-	L	A	-	T	L	F	M	-	-
Japanese quail- black	-	V	-	-	M	-	-	-	V	-	-	L	A	**K***	T	L	F	M	-	-

									110										120
Squirrel- E^{+}	L	E	V	G	A	L	A	T	P	A	T	V	V	Q	Q	L	D	N	V	I
Squirrel- E^{a}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Pig- MC1R*1	-	-	A	-	-	-	-	A	Q	-	A	-	-	-	-	-	-	-	-	M
Pig- MC1R*2	-	-	A	-	-	-	-	A	Q	-	A	-	-	-	-	-	-	-	-	M
Mouse- E^{+}	-	-	-	-	I	-	V	A	R	V	A	L	-	-	-	-	-	-	L	-
Mouse- E^{wo}	-	-	-	-	I	-	V	A	R	V	A	L	-	-	-	-	-	-	L	-
Mouse-E	-	-	-	-	I	-	V	A	R	V	A	L	-	-	-	-	-	-	L	-
Cattle-E^{+}	-	-	A	-	V	-	V	-	Q	-	A	-	-	-	-	-	-	-	-	-
Cattle-ED	-	-	A	-	V	-	V	-	Q	-	A	-	-	-	-	-	-	-	-	-
Jaguar- wt	-	-	A	-	T	-	-	G	R	-	A	-	-	-	-	-	-	D	-	-
Jaguar- mel	▲	▲	▲	▲	▲	T	-	G	R	-	A	-	-	-	-	-	-	D	-	-
Jaguarundis- red	-	-	A	-	T	-	-	G	R	-	A	-	-	-	-	-	-	D	I	-
Jaguarundis- dark	▲	▲	A	-	T	-	-	G	R	-	A	-	-	-	-	-	-	D	I	-
Rabbit- wt	-	-	A	-	-	-	-	G	R	-	A	-	-	-	-	-	-	D	-	-
Rabbit- E^{D}	-	-	A	-	-	-	-	G	R	-	A	-	-	-	-	-	-	D	-	-
Bananaquit- Y	-	-	H	-	V	-	V	M	R	P	S	I	-	R	H	M	-	S	-	-
Bananaquit- M	-	-	H	-	V	-	V	M	R	P	S	I	-	R	H	M	-	S	-	-
Chicken- wt	M	-	H	-	V	-	V	I	R	-	S	I	-	R	H	M	-	-	-	-
Chicken- B	M	-	H	-	V	-	V	I	R	-	S	I	-	R	H	M	-	-	-	-
Japanese quail-wt	M	-	H	-	V	-	V	I	R	-	S	I	-	R	H	M	-	-	-	-
Japanese quail-bl	M	-	H	-	V	-	V	I	R	-	S	I	-	R	H	M	-	-	-	-

Table 3.8 shows that many mutations associated with melanism involve removal of glutamic acid from TM2 or ECL1. In most cases this is E91 but in the jaguar this is E99. The replacement of this acidic residue with lysine in the mouse, chicken and lemur is known to lead to constitutive activation (Robbins et al., 1993; Lu, Vage and Cone, 1998; Haitina et al., 2007). Models of the MC1R suggest that the negatively charged glutamic acid is part of an acidic domain that interacts with the positively charged arginine of the α-MSH (García-Borrón, Sánchez-Laorden and Jiménez-Cervantes, 2005). Removal of this acidic residue is thought to mimic ligand binding leaving the receptor constitutively active in the absence of the ligand (Haskell-Luevano et al., 1996; Lu, Vage and Cone, 1998).

Table 3.8. Alignment of amino acids 85-110 of the melanocortin-1 receptor

Dashes indicate consensus with the wildtype squirrel sequence. Amino acid sequences of the MC1R associated with melanism are shown for each species. GL Tamarin = golden lion tamarin, G headed L tamarin = golden-headed lion tamarin. The red line highlights E91 and the green line highlights E99.

Wildtype	85					90					95					100					105					
Squirrel	S	T	S	N	A	L	E	T	T	I	F	L	L	L	E	V	G	A	L	A	T	P	A	T	V	V
Jaguarundis	-	V	-	S	V	-	-	-	A	V	M	-	-	-	-	A	-	T	-	-	G	R	-	A	-	-
Jaguar	-	V	-	S	V	-	-	-	A	V	M	-	-	T	-	A	-	T	-	-	G	R	-	A	-	-
GL Tamarin	-	G	-	-	M	-	-	-	A	V	I	-	-	-	-	A	-	V	-	-	-	R	-	S	-	-
Rabbit	-	V	-	S	V	-	-	-	A	V	L	-	-	-	-	A	-	-	-	-	G	R	-	A	-	-
Mouse	-	V	-	I	V	-	-	-	-	-	I	-	-	-	-	-	-	I	-	V	A	R	V	A	L	-
Chicken	-	V	-	-	L	A	-	-	L	F	M	-	-	M	-	H	-	V	-	V	I	R	-	S	I	-
Quail	-	V	-	-	L	A	-	-	L	F	M	-	-	M	-	H	-	V	-	V	I	R	-	S	I	-
Bananaquit	-	I	-	-	L	A	-	M	L	F	M	-	-	-	-	H	-	V	-	V	M	R	P	S	I	-
Melanic																										
Squirrel	-	-	■	■	■	■	■	■	■	■	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Jaguarundis	-	V	-	S	V	-	■	■	■	■	■	■	■	■	-	A	-	T	-	-	G	R	-	A	-	-
Jaguar	-	V	-	S	V	-	-	-	A	V	M	-	-	T	■	■	■	■	■	-	G	R	-	A	-	-
G headed L Tamarin	-	G	-	-	M	-	■	■	■	■	■	■	■	■	-	A	-	V	-	-	-	R	-	S	-	-
Rabbit	-	V	-	S	V	-	■	■	A	V	L	-	-	-	-	A	-	-	-	-	G	R	-	A	-	-
Mouse	-	V	-	I	V	-	K	-	-	-	I	-	-	-	-	-	-	I	-	V	A	R	V	A	L	-
Chicken	-	V	-	-	L	A	K	-	L	F	M	-	-	M	-	H	-	V	-	V	I	R	-	S	I	-
Quail	-	V	-	-	L	A	K	-	L	F	M	-	-	M	-	H	-	V	-	V	I	R	-	S	I	-
Bananaquit	-	I	-	-	L	A	K	M	L	F	M	-	-	-	-	H	-	V	-	V	M	R	P	S	I	-

3.3.2 The origin of the *MC1RΔ24* allele of the grey squirrel

The results presented here show that melanic grey squirrels of North America and Britain have an identical 24 bp deletion. This finding strongly supports the conclusion that the presence of melanic grey squirrels in Britain is the result of one or a few introductions from North America and not the result of a new mutation. This is consistent with reports by Shorten (1954) that approximately twelve melanic grey squirrels were introduced to Woburn some time after the introduction of wildtype grey squirrels. Squirrels from both Britain and North America showed the same phenotypic and corresponding genotypic variation where wildtype squirrels were grey and homozygous for the *MC1R-wt* allele, brown-black squirrels were heterozygous for the *MC1R-wt* and *MC1R-Δ24* allele and jet black squirrels were homozygous for the *MC1R-Δ24* allele (table 3.2).

3.3.3 Alleles of the *melanocortin-1 receptor* of the grey squirrel

Five alleles of the *MC1R* in the grey squirrel were detected, four being wildtype alleles. The number of alleles detected in a relatively small sample was unexpected. These alleles were well distributed across the different regions of Britain and North America suggesting genetic diversity and a heterogeneous population in both countries (see table 3.4 and figure 3.21). This diversity suggests that the population of grey squirrels in Britain originated from a genetically diverse founding population. The squirrels of Cumbria, however, were all homozygotes for allele one but there were only four samples from this location, all obtained from the same source and likely to be closely related. Brown-black squirrels also had all four of the wildtype alleles suggesting that all phenotypes interbreed freely. The genetic diversity present in the British population could partly explain why the grey squirrel has not apparently suffered from inbreeding but has been a successful invasive species in Britain. Repeated

introductions from different stocks from North America could have contributed to hybrid vigour in this translocated population and this may have increased their chances of survival even from a small founding population. Further genetic analysis would be needed to confirm these speculations.

In contrast to the *MC1R-wt*, sequences of the *MC1R-Δ24* allele showed only one haplotype, even though samples were tested from diverse locations in North America and Britain. This result shows an interesting lack of diversity in the *MC1R-Δ24* allele. This could be due to the small sample size from North America. The British samples were all taken from the same area and are likely to have originated from a small number of individuals originally introduced from North America. Therefore, it would not be surprising to find that all *MC1R-Δ24* alleles were the same. Given that the *MC1R-Δ24* is derived from the wildtype allele, it clearly has a more recent origin and is therefore less likely to have polymorphisms. Intriguingly, as well as being less varied than the wildtype allele, the *MC1R-Δ24* is as close a match to the fox squirrel *MC1R* as to the *MC1R-wt* with 97% identity. Interestingly, the SNPs unique to the *MC1R-Δ24* allele were also found in the *MC1R* of the fox squirrel. In fact, apart from the deletion, both alleles of the fox squirrel *MC1R* showed only two bp difference from the *MC1RΔ24* allele, as shown in table 3.8 compared to three bp or more to the *MC1R-wt.* The similarities between the MC1RΔ24 and the MC1R of the fox squirrel could suggest that a hybridisation event has taken place between the two species, however, this seems extremely unlikely given that the genetic basis of melanism is not the same between the two species. One explanation for this is that there are more alleles of the *MC1R-wt* and *MC1RΔ24* allele that have not as yet been detected. The most likely explanation is that the deletion allele is old. A wider sampling of both British and North American samples would be required to

confirm this. Overall, the *MC1R* sequences revealed a very close match (99.7% identity) between the grey and fox squirrel. This suggests that the two species are closely related and confirms phylogenetic and protein analysis work carried out which indicated that the two species diverged relatively recently at 9.8-14.4 million years ago (Ellis and Maxson, 1980). In contrast, the *MC1R* sequence of the red squirrel was 97% similar to the grey squirrel. This greater difference is consistent with phylogenetic analysis which estimates that New World and Old World Sciurus diverged 22.4-26.6 million years ago (Ellis and Maxson, 1980; Moncrief, Lack and Van Den Bussche, Ronald A, 2010). Figure 3.24 shows the phylogenetic relationship between the grey, fox and red squirrel (Herron, Castoe and Parkinson, 2004).

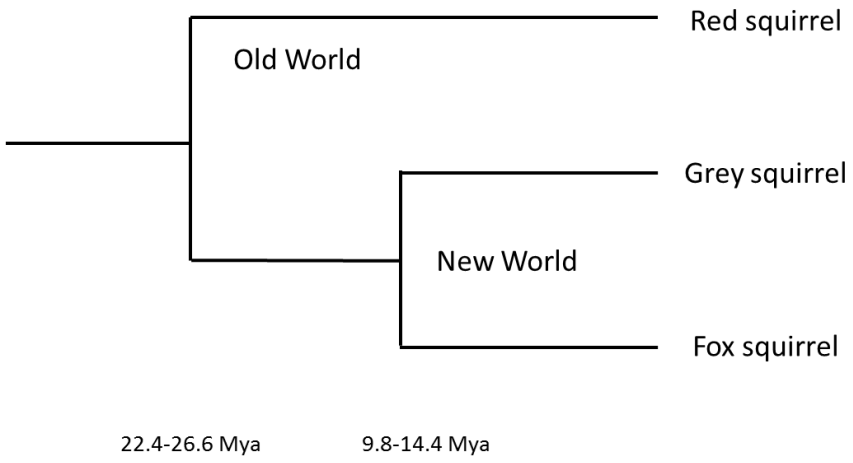


Figure 3.24. Phylogenetic relationships between the grey (*Sciurus carolinensis*), fox (*Sciurus niger*) and red (*Sciurus vulgaris*) squirrels (Herron, Castoe and Parkinson, 2004)

Mya = million years ago.

Apart from the deletion, there were no non-synonymous substitutions identified in the grey squirrel *MC1R*, however, there were seven synonymous mutations detected. This suggests that the MC1R is under purifying selection. Genes are considered to be under purifying selection where the ratio of the rate of non-synonymous to synonymous substitutions is below one. A larger sample size would be needed to confirm that this was the case here.

3.3.4 Alleles of the *melanocortin-1 receptor* of red and fox squirrel

There were a number of amino acid changes between the MC1R of the grey, fox and red squirrel as represented in figure 3.25. In the fox squirrel, the two alleles identified did not result in any identifiable phenotypic changes. The phenotypes of grey and fox squirrels are very similar, having banded hairs in the wildtype and an overall grizzled appearance. The red squirrel, however, has no banding and an overall reddish appearance in the wildtype. The

S10C and R233C differences could be changing the overall stability of the protein with changes to disulphide bridges maintaining the overall structure of the receptor. These cysteines are not known to be involved in disulphide bridges, but this does not rule out the possibility their involvement here.

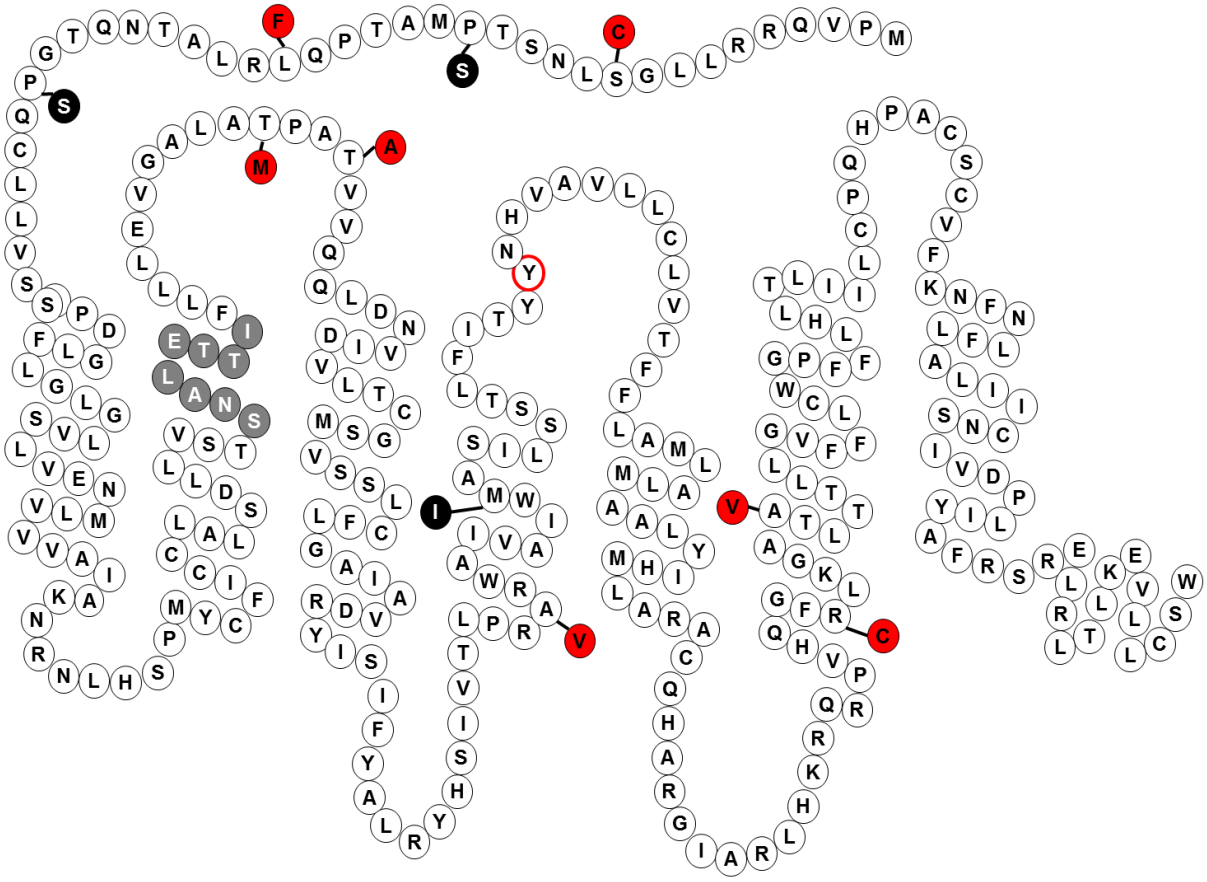


Figure 3.25. Schematic representation of the melanocortin-1 receptor showing amino acids of the wild-type grey, fox and red squirrels

Dark grey circles with white lettering indicate the 8 amino acids deleted in the MC1R-Δ24 allele. Black circles and red circles indicate amino acids in the fox and red squirrels, respectively that differ from those of the grey squirrel. The circle with red outline is the amino acid deleted in allele 2 of the red squirrel MC1R.

3.3.5 The genetic basis of melanism in the fox and red squirrel

These results show no association between variation in the *MC1R* gene and melanism in either the fox or red squirrel. In cases where melanism is associated with variations in this

gene, phenotypic differences are often discrete and relatively large, as demonstrated in the grey squirrel, bananaquit, chicken, mouse, jaguar, and pig (Robbins et al., 1993; Takeuchi et al., 1996; Kijas et al., 1998; Theron et al., 2001; Eizirik et al., 2003; McRobie, Thomas and Kelly, 2009). In both the fox and red squirrels, however, there are not such clear distinctions between phenotypes and the colour differences are more subtle, presenting a more continuous spectrum of colour variation. A similar spectrum of variation is also observed in the gopher (Wlasiuk and Nachman, 2007), three mustelid lineages (Hosoda et al., 2005), Old World leaf warblers (MacDougall-Shackleton et al., 2003), and the blue-crowned manakin (Cheviron, Hackett and Brumfield, 2006). In all of these cases, where there is a wide spectrum of colour morphs, melanism is not associated with variations in the *MC1R*. These findings suggest that cases where melanism is graduated across a species, genes other than the *MC1R* may be responsible. The *MC1R* gene is well characterized in a wide variety of species and the number of cases reporting the association of the *MC1R* to melanism indicates that this is a good candidate gene (Hoekstra et al., 2006). However, it is likely that there is an ascertainment bias, where positive results are more likely to be reported in this intronless gene, which is relatively easy to sequence and analyse (Mundy, 2005). Other key genes involved in pigmentation are more complex, for example, *ASIP*, which has 3 coding exons (Abdel-Malek et al., 2001) and is considerably harder to work with and therefore less likely to be reported. A number of other studies investigating the *MC1R* have highlighted the complex nature of the genetics of pigmentation. For example, large deletions in the first extracellular region of the receptor are associated with melanism in the jaguar (15 bp), jaguarundi (24 bp) (Eizirik et al., 2003), and grey squirrel (24 bp) (McRobie, Thomas and Kelly, 2009). In contrast, deletions almost identical to these are not associated with melanism in the wolverine (15 bp), stone marten (28 bp), four species of martens (45 bp) (Hosoda et al.,

2005), and the gopher (21 bp) (Wlasiuk and Nachman, 2007). Further analysis of protein expression and function would be necessary to elucidate the effect of these deletions.

3.4 Conclusion

There are three distinct phenotypes of the grey squirrel; the wildtype, the brown-black and the jet black. Melanism in the grey squirrel is associated with a 24 bp deletion in the *MC1R* gene, encoding a receptor with eight deleted amino acids at positions 87–94 (SNALETTI). The wildtype grey phenotype is homozygous for the wildtype *MC1R-wt* allele, the jet black phenotype is homozygous for the *MC1RΔ24* allele and the brown-black phenotype is heterozygous. The *MC1RΔ24* allele is incompletely dominant to the wildtype allele. The melanic grey squirrel, found in the British Isles, originated from one or more introductions of melanic grey squirrels from North America. These findings led to further research into the function of the MC1R-wt compared to MC1RΔ24. It was predicted that the MC1RΔ24 receptor would be either hyperactive or constitutively active (see figure 1.22). Melanism is not associated with variations of the *MC1R* gene in the fox or red squirrels. These findings led to further research exploring the variations in the candidate gene *ASIP* and associations with melanism.

Chapter Four

The Agouti Signalling Protein of the Grey, Fox and Red Squirrel

4.1 Introduction

Results presented in the previous chapter show that melanism in the fox and red squirrel is not associated with variation in the *MC1R* gene. Variations in the *ASIP* gene are associated with melanism in many mammal species including mice (Steiner, Weber and Hoekstra, 2007), foxes (Våge et al., 1997a) and rabbits (Fontanesi et al., 2010). This gene was therefore the next candidate for scrutiny in the investigation of the genetic basis of melanism in the fox and red squirrels. The hypothesis was that variations in the ASIP gene of the fox and red squirrels are associated with melanism.

The specific methods used in this section include DNA amplification by PCR, sequencing of PCR products and DNA sequencing. Sequences were aligned using BLASTN and Clustal Omega.

4.2 Results

4.2.1 The *agouti signalling protein* gene sequence from the grey, fox and red squirrel

The *ASIP* gene was amplified using genomic DNA from the grey, fox and red squirrel. Exon 2, the first coding exon, was amplified using the primers ASP2_13LF1 and ASP3sqR11. A primer map showing relative positions of the primers is shown in figure 2.3. This PCR product gave a band of approximately 3000 bp and figure 4.1 shows a typical gel picture for this reaction.

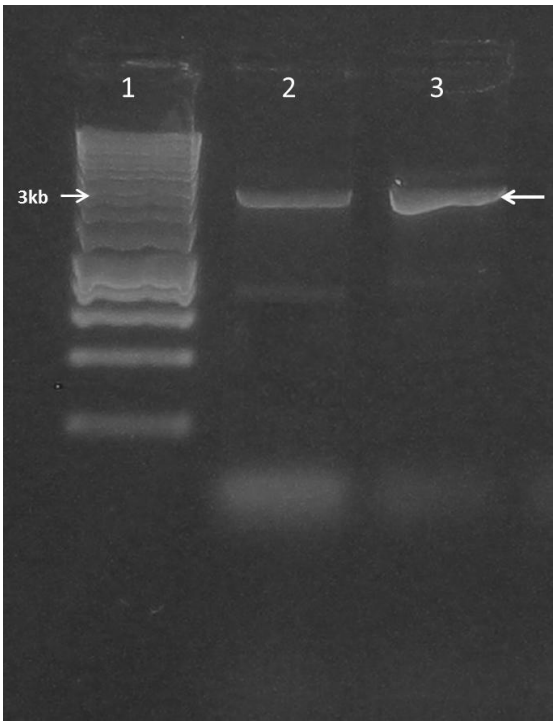


Figure 4.1. PCR amplification of exon 2 of the *agouti signalling protein* gene from the grey squirrel (*Sciurus carolinensis*)

Representative agarose gel electrophoresis (1.3%) of PCR amplification products visualised with ethidium bromide and illuminated with UV light. Lane 1, Hyperladder 1 (Bioline) DNA ladder, lanes 2 and 3 show a 3 kilo base (kb) product of exon 2 of the ASIP gene of the grey squirrel amplified with the forward primer ASP2_13LF1 and reverse primer ASP3sqr11. The arrow indicates the band of interest.

A PCR product from this reaction, using grey squirrel DNA, was purified and sequenced with the ASP2_13LF1 primer and figure 4.2 shows a typical chromatogram from sequencing this product.

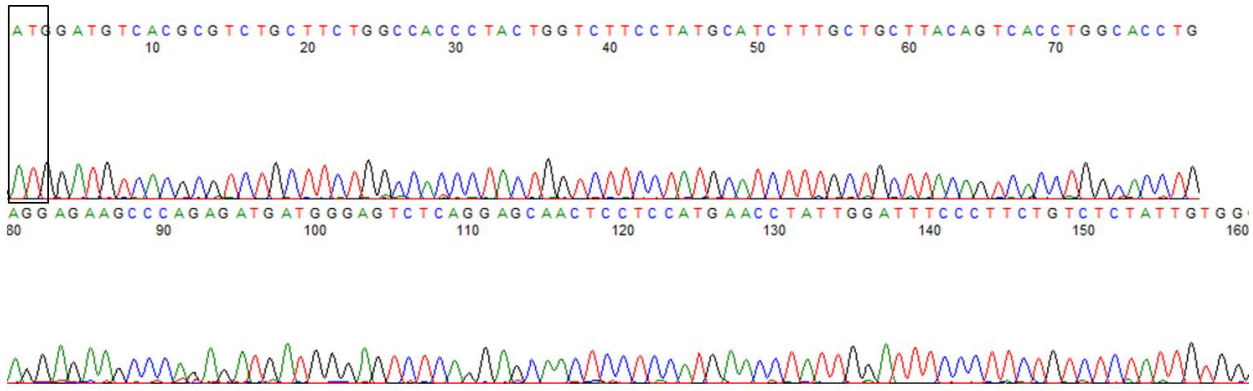


Figure 4.2. Chromatogram of exon 2 of the *agouti signalling protein* gene from the grey squirrel (*Sciurus carolinensis*)

Chromatogram obtained from sequencing an ASP_13LF1 and ASP3sqr11 PCR product using ASP2_13LF1 to sequence. This is the complete protein coding sequence of exon 2 (160 base pairs) of the ASIP gene. Numbers indicate base pairs. The box indicates the start codon.

Exon 3 was amplified with the primers ASPsqintron2to3F1 and ASPintron3to4R1 using grey squirrel DNA. This PCR product gave a band of approximately 200 bp. Figure 4.3 shows a typical gel for this reaction.

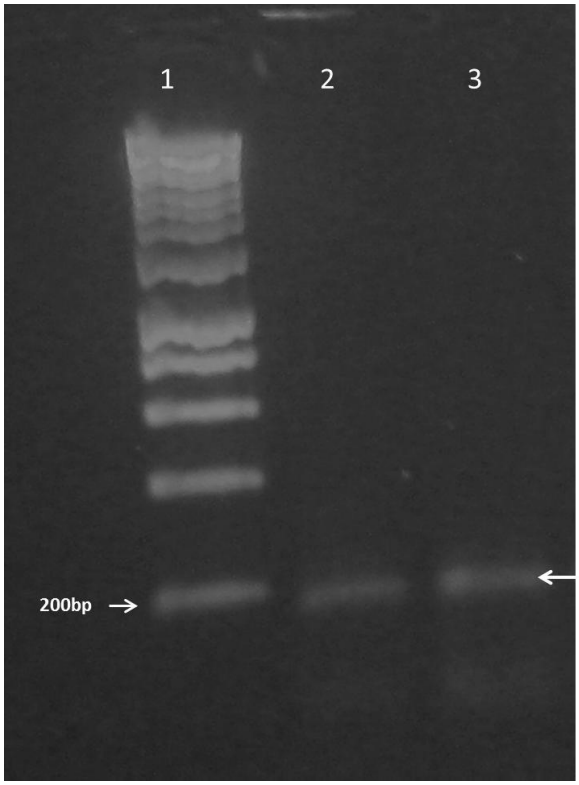


Figure 4.3. PCR amplification of exon 3 of the *agouti signalling protein* gene from the grey squirrel (*Sciurus carolinensis*)

Representative agarose gel electrophoresis (1.3%) of PCR amplification products visualised with ethidium bromide and illuminated with UV light. Lane 1, Hyperladder 1 (Bioline) DNA ladder; lanes 2 and 3 show a 200 bp product of exon 3 of the ASIP gene of the grey squirrel amplified with the forward primer ASPsqintron2to3F1 and reverse primer ASPintron3to4R1. The arrow indicates the band of interest.

The PCR product from this reaction was purified and sequenced with the ASPsqintron2to3F1 primer. Figure 4.4 shows a typical chromatogram obtained from sequencing this product.

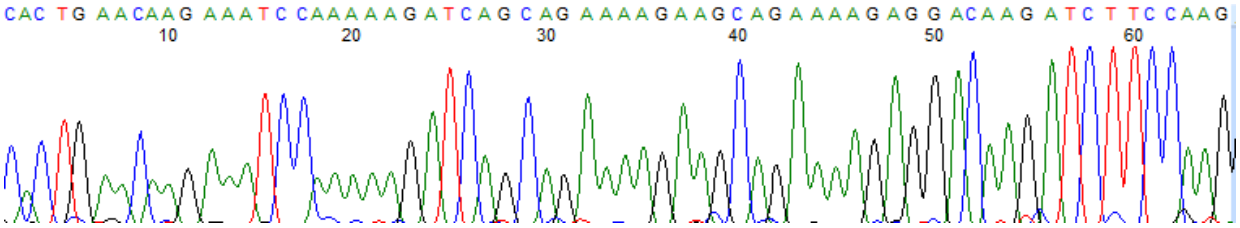


Figure 4.4. Chromatogram of exon 3 of the *agouti signalling protein* gene from the grey squirrel (*Sciurus carolinensis*)

Chromatogram obtained from sequencing an ASPsqintron2to3F1 and ASPintron3to4R1PCR product using ASPsqintron2to3F1 to sequence. This is the complete protein coding sequence of exon 3 (65 base pairs) of the ASIP gene.

Exon 4 was amplified with the primers ASP3to4sqF11 and ASP4sqrex30. This PCR gave a band of approximately 250 bp and figure 4.5 shows a typical gel picture for this reaction for the grey squirrel.

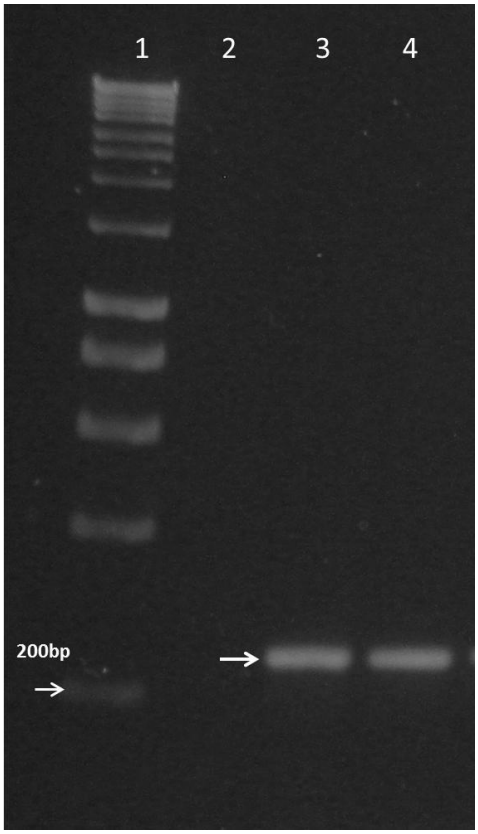


Figure 4.5. PCR amplification of exon 4 of the *agouti signalling protein* gene from the grey squirrel (*Sciurus carolinensis*)

Representative agarose gel electrophoresis (1.3%) of PCR amplification products visualised with ethidium bromide and illuminated with UV light. Lane 1, Hyperladder 1 (Bioline) DNA ladder; lane 2, negative control; lanes 3 and 4 show a 250 bp product of exon 4 of the ASIP gene of the grey squirrel amplified with the forward primer ASP3to4sqF11 and reverse primer ASP4sqrex30. The arrow indicates the band of interest.

The PCR product from this reaction was purified and sequenced with both forward and reverse primers and a contiguous sequence was constructed, as shown in figure 4.6 where the part of the chromatogram underlined in green was obtained from sequencing with the product with the ASP4sqrex30 primer and the part underlined in blue was obtained from sequencing the same product with the ASP3to4sqF11 primer.

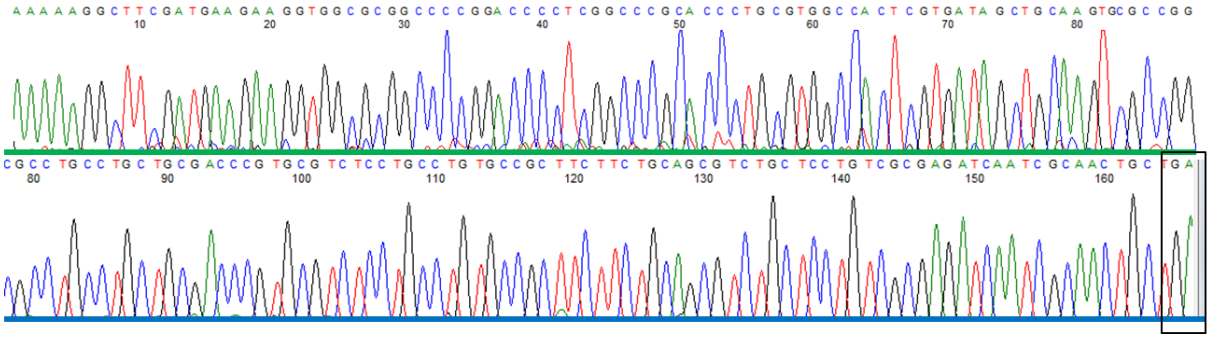


Figure 4.6. Chromatogram of the contiguous sequence of exon 4 of the *agouti signalling protein* gene of the grey squirrel (*Sciurus carolinensis*)

This sequence is obtained from sequencing an ASP3to4sqF11 and ASP4sqrex30 PCR product. The sequence underlined in green was obtained using ASP4rsqrex30 to sequence and the sequence underlined in blue was obtained using ASP3to4sqF11 to sequence. Exon 4 has a total of 177 base pairs. The box indicates the stop codon.

The sequences obtained were compared to the Genbank database using BLASTN to confirm that each PCR product was from the *ASIP* gene. The sequence for the grey squirrel has 92% identity with the *ASIP* of the thirteen-lined squirrel. Figure 4.7 shows an alignment of the *ASIP* of the thirteen-lined squirrel and the grey squirrel. Exons are highlighted in different colours. Exon 2 (blue) has 160 bp, exon 3 (pink) has 65 bp and exon 4 (green) has 177 bp giving a total of 402 bp which encodes a 133 amino acid protein.

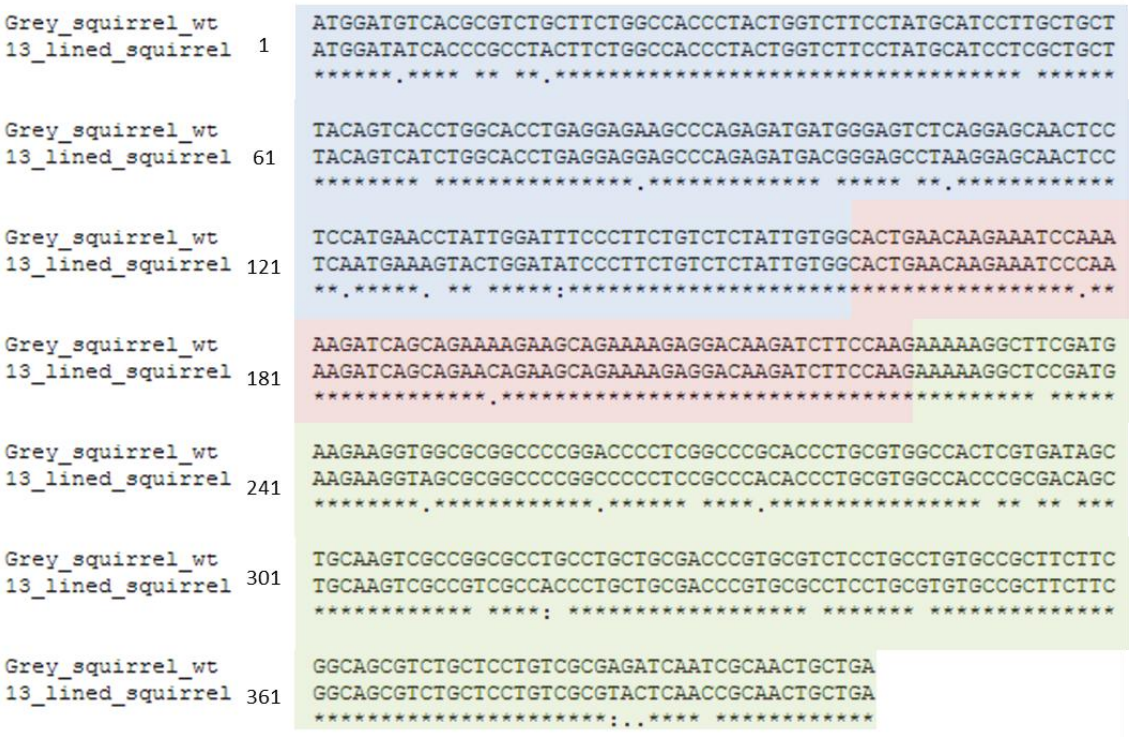


Figure 4.7. Alignment of the coding regions of the *agouti signalling protein* gene from the grey squirrel (*Sciurus carolinensis*) and the thirteen-lined squirrel (*Spermophilus tridecemlineatus*)

The sequence of the coding region of the ASIP gene of the thirteen-lined squirrel was obtained from Genbank accession number AAQQ01611796.1. Exon 2 is highlighted in blue, exon 3 in pink and exon 4 in green. Asterisks below the sequence indicate fully conserved base pairs, colons indicate strongly similar properties, full stop indicates weakly similar properties and spaces indicate no similarity between base pairs. Numbers refer to base pairs with the A of the start codon as one.

4.2.2 Variation between agouti signalling protein in the grey, fox and red squirrel

Having confirmed that the *ASIP* gene had been successfully amplified, the same primers were used to amplify the *ASIP* of the fox and red squirrels. The alignment of the gene from the wildtype grey, fox and red squirrels in figure 4.8 shows that there are no major differences between the *ASIP* sequences of the three squirrel species under consideration with 99% identity between each.

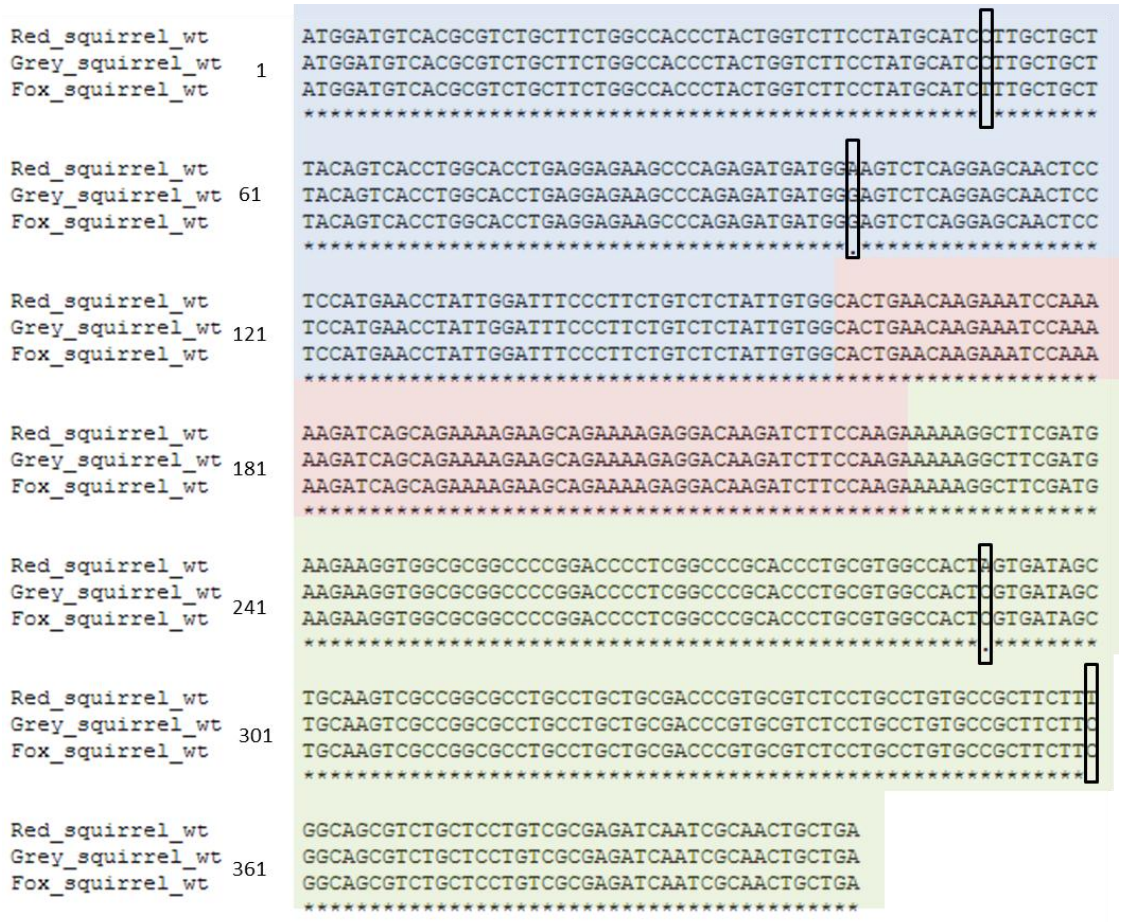


Figure 4.8. Alignment of the coding regions of the *agouti signalling protein* gene from the grey (*Sciurus carolinensis*), fox (*Sciurus niger*) and red (*Sciurus vulgaris*) squirrels

Exon 2 is highlighted in blue, exon 3 in pink and exon 4 in green. Asterisks below the sequence indicate fully conserved base pairs, dots indicate weakly similar properties and spaces indicate no similarity between base pairs. Numbers refer to base pairs with the A of the start codon as one. Differences between sequences are highlighted in the boxes.

Figure 4.9 shows an amino acid sequence alignment of the grey, fox and red ASIP. Each species has a unique ASIP sequence, however, there was 99% identity between each. There are two amino acid positions where variation is found between the species: the grey squirrel

has L18 and R98, the fox squirrel has F18 and R98 and the red squirrel has L18 and S98, as shown in figure 4.9

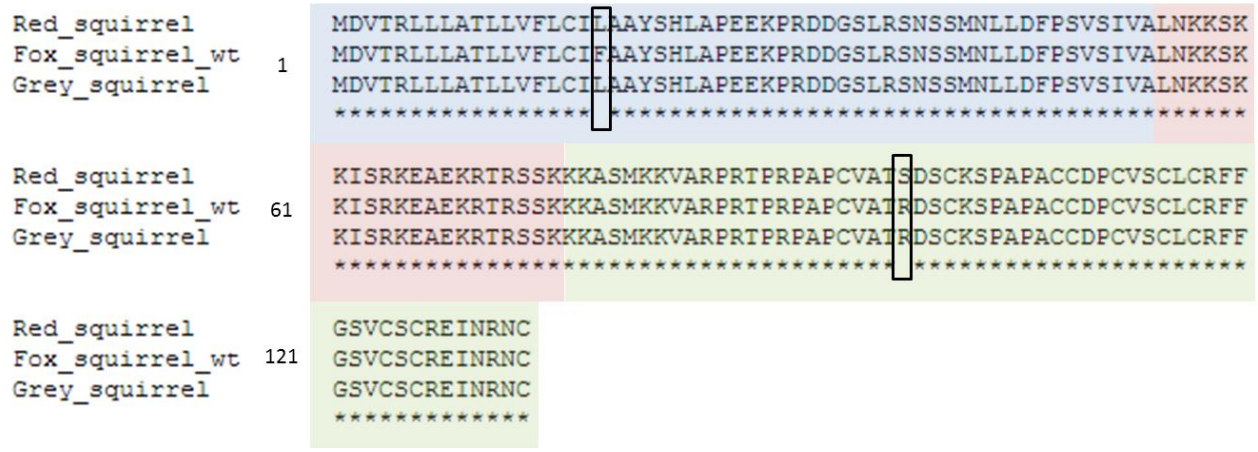


Figure 4.9. Alignment of amino acid sequences of agouti signalling protein from the red squirrel (*Sciurus vulgaris*), fox squirrel (*Sciurus niger*) and grey squirrel (*Sciurus carolinensis*)

Sequences shown are all from the wildtype of each species. The boxed areas highlight the amino acid differences between species. The boxed areas highlight amino acid differences between each case. Asterisks below the sequence indicate fully conserved amino acids and spaces indicate no similarity. Numbers refer to amino acids with the first amino acid M as one.

There were no intraspecific differences between sequences of the wildtype and melanic ASIP in either the grey or red squirrel (alignments not shown). However, there were differences found between the two morphs in the fox squirrel.

4.2.3 Melanism in the fox squirrel

Alignment of the *ASIP* sequences obtained from the wildtype and melanic fox squirrels show two bp differences corresponding to 253 C/G and 361 G/T respectively, as shown in figure 4.10.

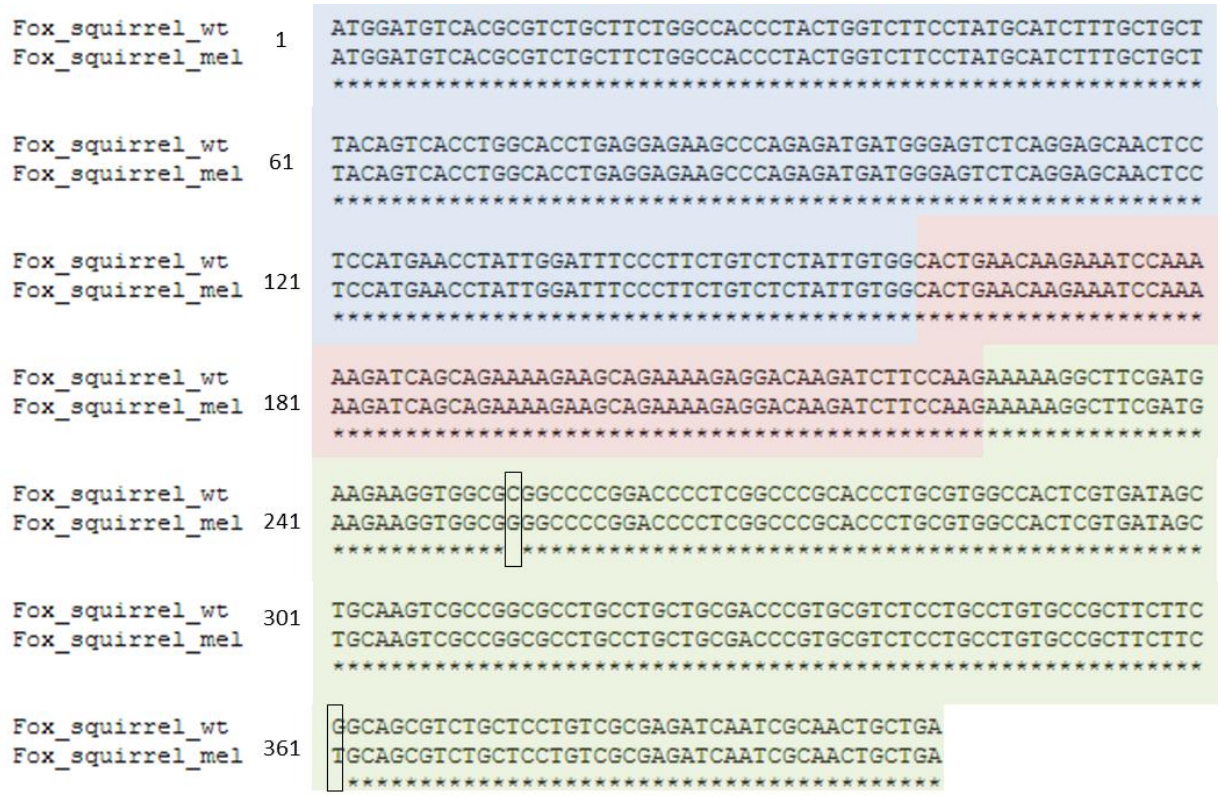


Figure 4.10. Alignment of the coding region of the *agouti signalling protein* gene from the wildtype and melanic fox squirrel (*Sciurus niger*)

Sequences of the coding region of the ASIP gene of the fox squirrel aligned with Clustal Omega. Exon 2 is highlighted in blue, exon 3 in pink and exon 4 in green. The boxed regions highlight the 253 C/G and 361 G/T mutations. Asterisks below the sequence indicate fully conserved base pairs and spaces indicate no similarity. Numbers refer to base pairs with the A of the start codon as one.

These SNPs are non-synonymous and amino acid differences between the two morphs are shown in figure 4.11. The differences correspond to R85G and G121C with the wildtype as the reference. Melanism is perfectly associated with these two amino acid changes and all samples were homozygous.

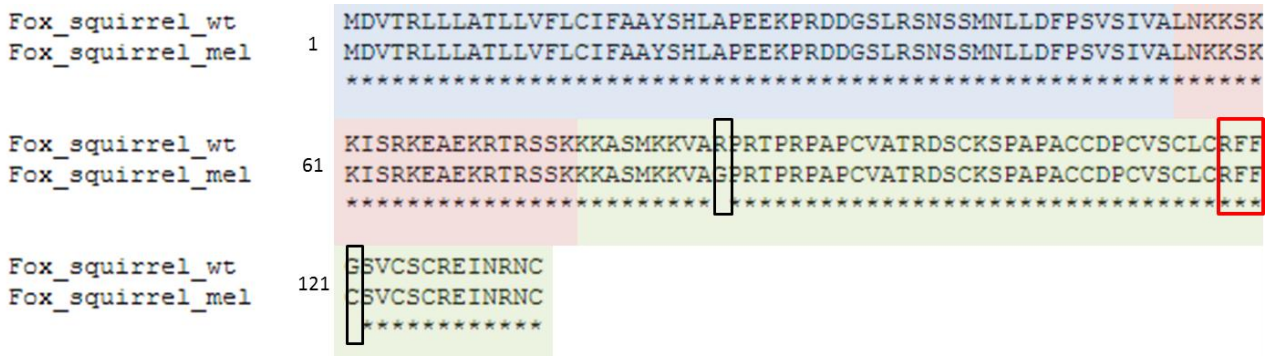


Figure 4.11. Alignment of amino acid sequences of agouti signalling protein from the wildtype and melanic fox squirrel (*Sciurus niger*)

The boxed areas highlight the R85G and G121C mutations associated with melanism. Highlighted with the red box is the highly conserved RFF sequence essential for ASIP function. Asterisks below the sequence indicate fully conserved amino acids and spaces indicate no similarity. Numbers refer to amino acids with the first amino acid M as one.

4.2.4 Melanism and the red squirrel

There was no difference found between the wildtype and melanic sequences of *ASIP* in the red squirrel. To investigate the *ASIP* gene further, splice donor and acceptor sites of the exons were examined to establish if there were any mutations associated with melanism. Figure 4.12 shows the splice sites on the chromatograms highlighted in blue. All splice sites were identical for all three species showing the conserved donor GT sequence following an exon and conserved acceptor AG sequence preceding an exon.

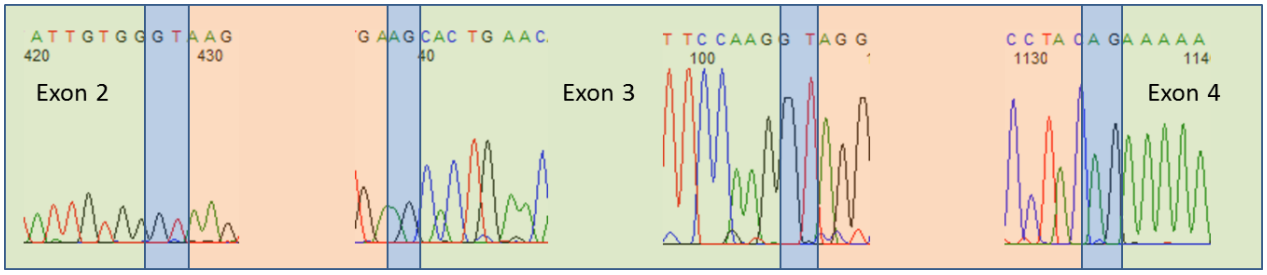


Figure 4.12. Splice donor and acceptor sites of the *agouti signalling protein* gene from the red squirrel (*Sciurus vulgaris*)

Green sections show parts of the exons two, three and four, and orange sections show the introns and the blue sections highlight the splice sites.

4.2.5 α-Melanocyte stimulating hormone of the grey squirrel

The α-MSH is a 13 residue peptide with a highly conserved sequence of SYSMEHFRWGKPV (Dores and Baron, 2011). Functional analyses of the MC1R in the present study use α-MSH purchased from Sigma with this sequence. To confirm that this is the same sequence found in the grey squirrel, the gene was amplified and sequenced here. The *POMC* gene was amplified using genomic DNA from the grey squirrel. The primers MSHF1 and MSHR9 were used initially and the PCR product from this reaction (product from lane 6 figure 4.13 A) was used to perform a further "nested" PCR using the primers MSHF1 and MSHR4. Nested PCR was required as the first reaction produced multiple bands. Figure 4.13 shows the PCR gel pictures obtained from these reactions.

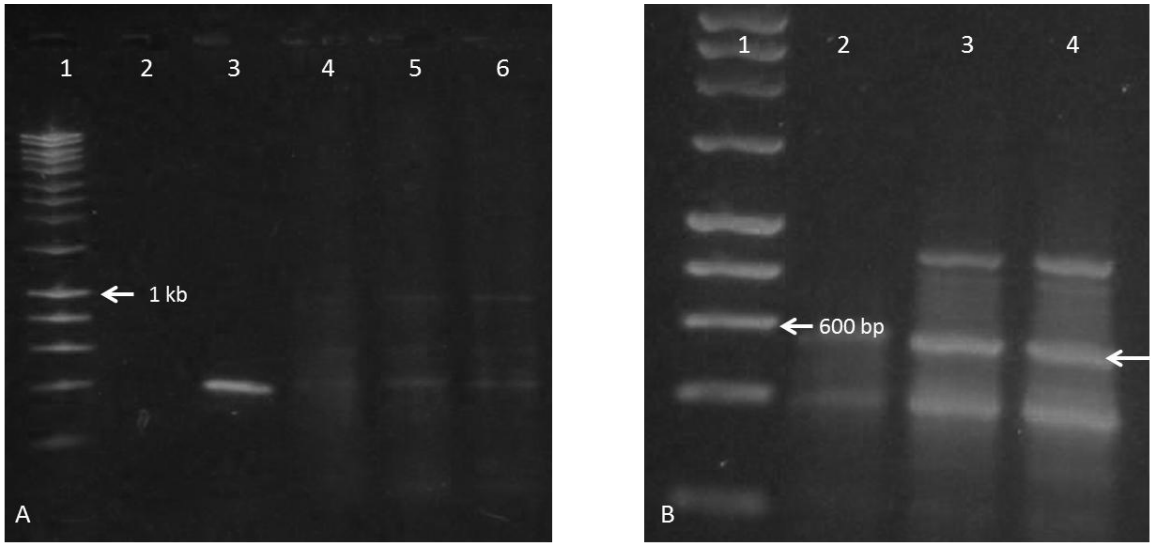


Figure 4.13. PCR amplification of the *α-melanocyte stimulating hormone* gene from the grey squirrel (*Sciurus carolinensis*)

***A**garose gel electrophoresis (1.3%) of PCR amplification products visualised with ethidium bromide and illuminated with UV light. A) Lane 1, Hyperladder 1 (Bioline) DNA ladder; lanes 2,3,4,5 and 6 show various products of the reaction using the primers MSHF1 and MSHR9. B) Lane 1, Hyperladder 1; lanes 2, 3 and 4 show multiple products of the nested PCR using the product of reaction A; lane 6 as template with the forward primer MSHF1 and reverse primer MSHR4. The arrow indicates the α-MSH. PCR and electrophoresis performed by Craig Simpson, Anglia Ruskin University.*

A band of approximately 600 bp was extracted from the agarose gel (indicated with an arrow in figure 4.13 B) and sequenced. The results showed that the sequence predicted to be that of the *POMC* gene encoding α MSH is identical to that of all other reported sequences from mammals including the human sequence (SYSMEHFRWGKPV) (figure 4.14).

```
Homo_sapiens        AGCTATAGCATGGAACATTTTCGCTGGGGCAAACCGGTG
Grey_squirrel       AGCTATAGCATGGAACATTTTCGCTGGGGCAAACCGGTG
                    ****************************************
```

Figure 4.14. Alignment of the *α- melanocyte stimulating hormone* gene from humans and the grey squirrel *(Sciurus carolinensis*)

The human sequence is from Genbank data base accession number NM_001035256.1.

4.3 Discussion

4.3.1 Melanism in the fox squirrel

The three protein coding exons of the *ASIP* gene were sequenced for the grey, fox and red squirrel. These sequences had 92% identity with that of the thirteen-lined squirrel from the Genbank database. There were no differences between wildtype and melanic *ASIP* sequences in the grey squirrel. However, melanism in the fox squirrel was found to be perfectly associated with two amino acid substitutions in ASIP. The differences correspond to R85G and G121C with the wildtype as reference. ASIP is made up of five structural domains, as shown in figure 4.15.

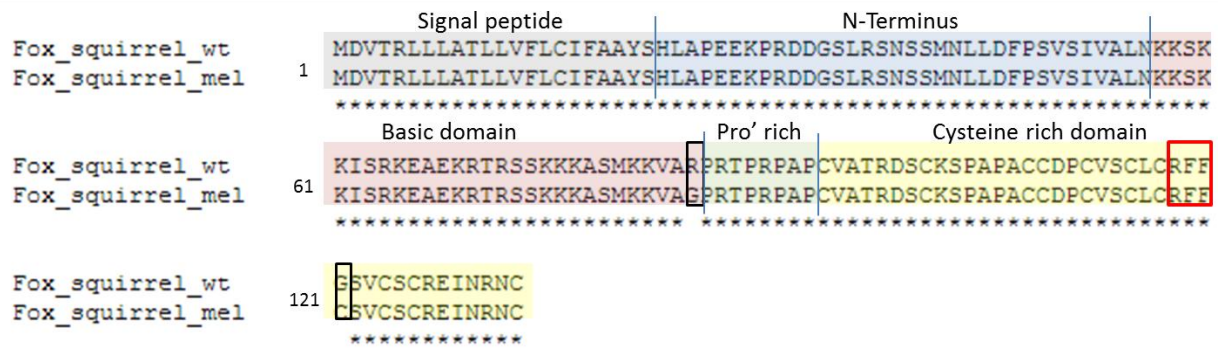


Figure 4.15. Amino acid alignment of agouti signalling protein from the wildtype and melanic fox squirrel (*Sciurus niger*)

Protein domains are indicated above the sequence with the signal peptide domain coloured in grey, the N-terminus domain in blue, the basic domain in pink, the proline-rich domain in green and the cysteine-rich domain in yellow. The amino acid differences between the morphs are highlighted in boxes and the highly conserved RFF motif is highlighted in the red box. (Pro'rich = proline-rich). Numbers refer to amino acids with the first amino acid, M as one.

The cleavable signal peptide is essential for entry into the secretory pathway (Perry et al., 1996). The N-terminus is involved in interactions with the accessory protein attractin

(Jackson et al., 2006). The basic domain is highly conserved and involved in activity *in vivo* but its precise role is not fully understood (Perry et al., 1996). The proline-rich domain is thought to provide a flexible hinge between the basic and the cysteine-rich domains (Miltenberger et al., 2002). The most critical and best understood region of ASIP is the cysteine-rich domain, which is thought to form a highly ordered structure stabilised by five disulphide bonds (Bolin et al., 1999). The positions of the ten cysteines involved are highly conserved and found in all functional ASIP proteins (Miltenberger et al., 2002). The disulphide bonds create an "inhibitor cystine-knot" motif which forms three constrained loops; the N-terminal loop, the active loop and the C-terminal loop shown in figure 4.16. The active loop carries the highly conserved RFF sequence which is essential for interaction with the MC1R (Tota et al., 1999).

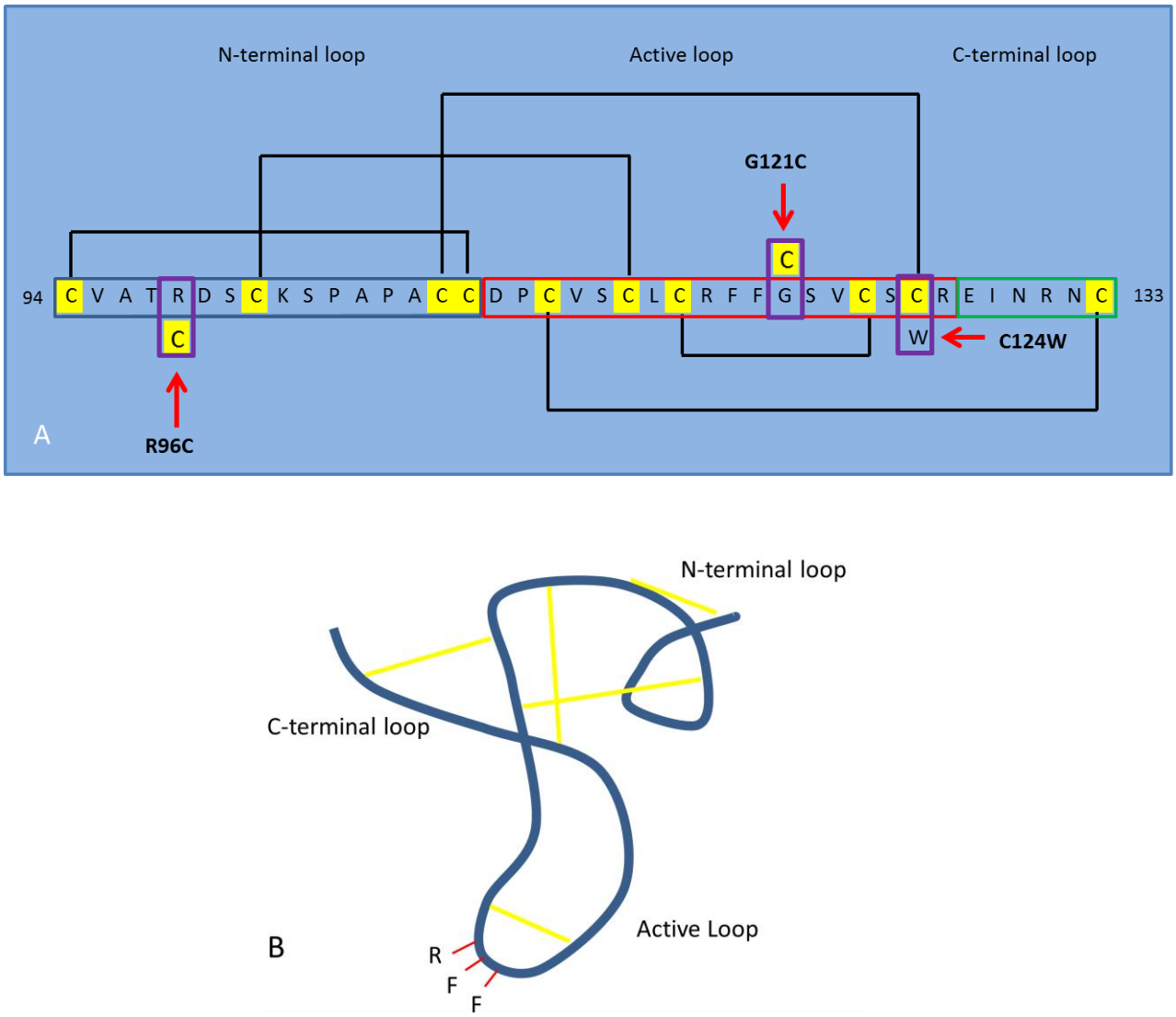


Figure 4.16. Cysteine-rich domain of the agouti signalling protein

A) Sequence of the cysteine-rich domain of the fox squirrel (Sciurus niger) showing the N-terminal loop boxed in blue, the active loop boxed in red and the C-terminal loop boxed in green. Disulphide bridges are shown as black lines connecting cysteines which are coloured yellow. The bridges are predicted from those of mouse and human ASIP (McNulty et al., 2005). Mutations in the dog (Canus lupus) (R96C), (Kerns et al., 2004) fox squirrel (G121C) and Asian Golden cat (Pardofelis temminckii) (C124W), (Schneider et al., 2012) are highlighted. B) Schematic representation of the cysteine-rich domain showing the constrained loops formed by the disulphide bridges. The RFF motif is highlighted.

4.3.2 Mutations associated with melanism and variation in agouti signalling protein

Mutations associated with melanism in ASIP are loss of function mutations and are normally recessive (Nachman, Hoekstra and D'Agostino, 2003). There are many more ways to "break" a protein than make it work more efficiently, and loss of function mutations are found throughout the coding and regulatory regions of ASIP as discussed in the introduction. Appendix 3 shows an alignment of ASIP sequences with mutations associated with melanism. Eight out of twelve of these mutations involve either complete deletion of the gene or premature stop codons leading to a truncated protein which either leads to the complete absence or a completely non-functional protein. Only four of these mutations leading to melanism are caused by substitutions and of these, two involve either gain or loss of cysteines in the cysteine-rich domain.

4.3.3 Domains of agouti signalling protein

The R85G mutation found in the melanic fox squirrel falls in the basic domain and it seems likely that the loss of the basic Arg residue in the melanic morph is biologically significant. Extensive studies on mice with artificially induced mutations show that the loss of a basic residue from this domain is associated with a darkened phenotype (Miltenberger et al., 2002). It has been suggested by Miltenberger et al (2002) that the loss of an overall net charge in this domain may be causing the loss of function in these mutants. This loss of charge could be affecting interactions with the MC1R or the accessory protein attractin or it could be affecting protein trafficking or processing (Miltenberger et al., 2002). Mutations in the basic domain were found to have relatively mild effects in the mice. In contrast, mutations which caused the most severe effects were located in the cysteine-rich domain (Miltenberger et al., 2002). It has been found that the loss of cysteines causes a general loss of function in the protein leading to a darker phenotype in mutant mice (Miltenberger et al., 2002). An example of the loss of a cysteine leading to melanism is the naturally occurring C124W mutation in the Asian golden cat (Schneider et al., 2012). The *addition* of a cysteine seems to have an equally dramatic effect leading to the completely black phenotype in the R96C mutation found in German Shepherd dogs (Kerns et al., 2004). The G121C substitution in the fox squirrel is likely to be significant here as this extra cysteine also falls in the cysteine-rich domain of the protein, as shown in figure 4.15 and 16. Figure 4.16 shows the disulphide bridges predicted to be formed in the fox squirrel ASIP based on the cysteine bridges found in the mouse and human protein (McNulty et al., 2005). Figure 4.16 B shows the gain and loss of cysteines in the R96C, G121C and C124W mutations of the dog (Kerns et al., 2004), fox squirrel and Asian golden cat (Schneider et al., 2012) respectively. It seems likely that extra cysteines could form aberrant disulphide bridges and in this way alter the highly ordered and

constrained structure and function of the protein (McNulty et al., 2005). Figure 4.15 and 4.16 show the highly conserved RFF sequence of the constrained active loop highlighted. It has been suggested by McNulty et al., (2005) that the bridges in the active loop facilitate the presentation of the RFF sequence to the MC1R binding pocket where it is thought to interact with acidic and aromatic residues. The extra cysteine in the melanic fox squirrel is adjacent to this RFF sequence and it seems likely that an altered disulphide bridge here would have a particularly detrimental impact on ASIP function by altering the stabilisation of the essential RFF region of the protein.

4.3.4 Possible epistatic interactions with agouti signalling protein in the fox squirrel

It is unusual to have two mutations that may both have an effect on protein function as appears to be the case here in the fox squirrel. It might be the case that both mutations have disruptive effects on ASIP and both cause melanism to varying degrees. It is possible that an individual with one of the mutations but not the other has only a moderately darker phenotype and an individual with both mutations has a considerably darker phenotype. This would be consistent with the spectrum of colour morphs found in the fox squirrel ranging from grizzled russet-orange through various shades of grey to black. This spectrum is in contrast to the three discrete morphs of grey, brown-black and jet black found in the grey squirrel. There is also a possibility of epistatic interactions between the MC1R and ASIP as is the case with beach mice (Steiner, Weber and Hoekstra, 2007) and horses (Rieder et al., 2001). It must be noted that the number of samples available was small with only four wildtype fox squirrels and five melanic fox squirrels and therefore no clear conclusions can be reached at this stage. The two mutations presented here are indeed intriguing; however, the

lack of available samples and the constraints of time mean that a fuller investigation is beyond the scope of this study. Further studies in this area will form part of future work to be carried out into the association between melanism and these mutations in *ASIP* of the fox squirrel.

4.3.5 Melanism and the red squirrel

These results show no association between melanism and variations in *ASIP* coding regions or splice donor or acceptor sites in the red squirrel. The genetic basis of melanism in this species could be associated with regulatory regions of either the *MC1R* or *ASIP*. Mutations in regulatory regions have been found in sheep (Norris and Whan, 2008) , the Boxer dog (Ciampolini et al., 2013) and possibly Mangalitza pigs (Drögemüller et al., 2006) and cattle (Girardot et al., 2005) as described in the introduction. Alternatively, the genetics of melanism in the red squirrel may involve a different gene or may be polygenic with many different genetic factors contributing to the overall phenotype. Melanism in the three species of squirrel in this study gives a good example of convergent evolution where different genetic mechanisms lead to similar phenotypes with presumably the same ecological functions.

4.3.6 Alpha-melanocyte stimulating hormone of the grey squirrel

The α-MSH sequence of the grey squirrel is identical to that of all other α-MSH sequences deposited in Genbank at present (NCBI, 2013). The sequence includes the conserved HFRW sequence which is considered to be essential to its function as an agonist to the MC1R. The amino acid sequences of ASIP and α-MSH from the grey squirrel are used in the next part of

the study to investigate the ligand-receptor interactions, structure and function of the MC1R in the grey squirrel.

4.4 Conclusion

Melanism is associated with variations of the *ASIP* gene in the fox squirrel, however, the sample size is small and further studies are required to confirm this finding. There is no association between melanism and variations of the *ASIP* gene in the red squirrel. The sequence of α-MSH of the grey squirrel contains the highly conserved HFRW motif essential for activation of the MC1R.

Chapter Five

Structural Studies of the Melanocortin-1 Receptor of the Grey Squirrel

5.1 Introduction

At present, little is known about the precise structure of melanocortin receptors. Solving GPCR structures is notoriously difficult but there are many computational methods available that can be used to predict molecular structures from amino acid sequences. A combination of homology modelling and experimental data can provide structural information about these complex and flexible molecules. The aim of this part of the study was to use the available methods and knowledge to predict the structures and ligand interactions of the MC1R-wt and MC1RΔ24 in order to understand how the 24 bp deletion affects the structure and function of the receptor. The hypothesis here was that sequence variation between the MC1R-wt and MC1RΔ24 leads to differences in structure of the MC1R between each case.

The methods used in this section included the use of Phyre2, I-TASSER and Pepfold web servers for prediction of the three dimensional structure of molecules under consideration and the use of Molsoft and Hex 8.0 to visualise pdb files. Pdb files were analysed with the programmes Socket, Fpocket and Pepsite and molecular docking was carried out using *CLC Discovery*.

5.2 Results

5.2.1 Melanocortin-1 receptor homology model

Computer predictions of the structures of the MC1R-wt and MC1RΔ24 receptors were obtained from Phyre2 and I-TASSER web servers. Figures 5.1, 5.2 and 5.3 show the PSI-BLAST alignments from each server. PSI-BLAST uses a powerful algorithm to iteratively search the protein sequence database of approximately 100 million sequences. This method of searching allows remote homologues to be detected where less than 30% sequence identity can be used to predict protein structure (Beuming and Sherman, 2012). Figure 5.1 gives a general overview of the PSI-BLAST alignment from Phyre2, purposely lacking in fine detail, in order to get an overall sense of regions and features that are conserved or varied. It shows that, in general, transmembrane helices are more conserved than loops and that the N-terminal regions are particularly variable in length.

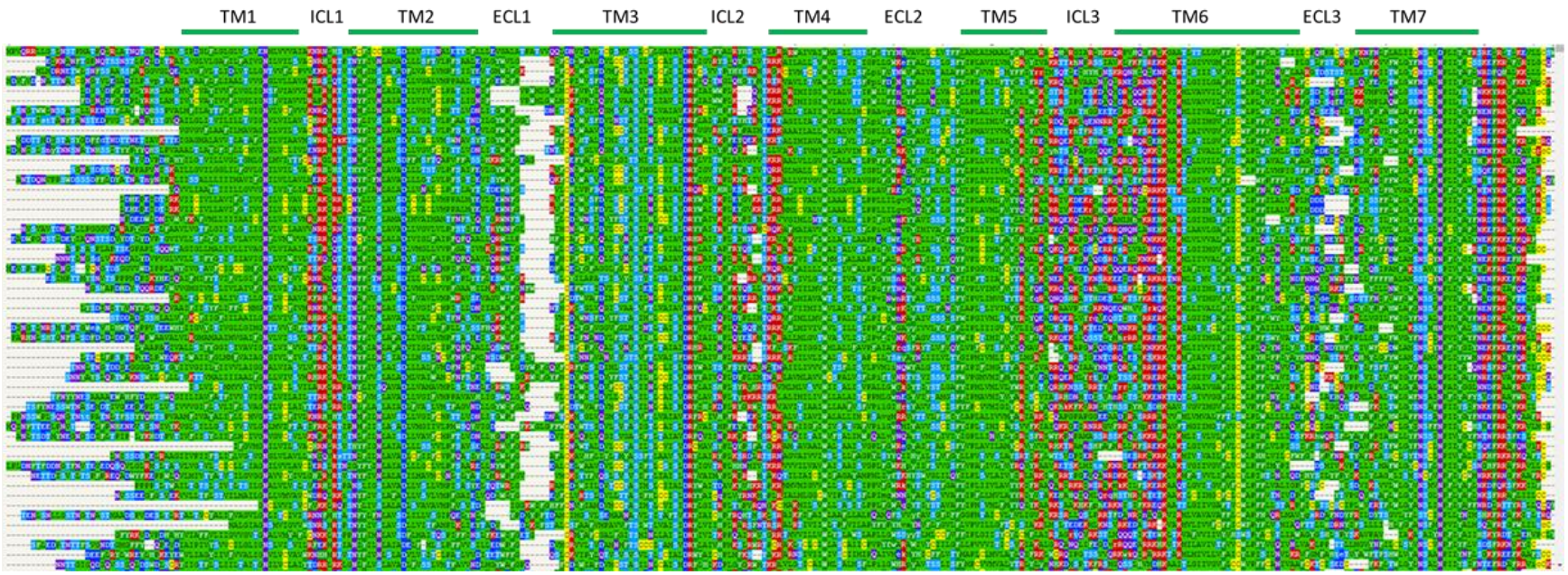


Figure 5.1. Overview of the PSI-BLAST sequence alignment obtained and adapted from Phyre2

The general overview allows the whole sequence to be viewed and highlights the overall conserved nature of the sequence. The MC1R-wt sequence is on the top line. TM = transmembrane domains, ICL = intracellular loops and ECL = extracellular loops. Colours: yellow = cysteine, red = lysine or arginine, purple = asparagine or glutamine, dark blue = glutamic acid or aspartic acid, light blue = threonine or serine. All other amino acids are green.

Figure 5.2 shows a more detailed view of the same PSI-BLAST alignment from Phyre2 as figure 5.1, showing only the top ten matches with the MC1R-wt. This alignment shows that the MC1R has the longest ECL1 compared to other sequences. The following residues are conserved in all ten sequences in figure 5.2; N53, D81, R139 (of the DRY motif), W166, F247, L283, S288, N287, P282 and Y295. The cysteine (114) that is conserved in all other sequences is replaced with aspartic acid in the MC1R-wt. This acidic residue forms an important part of an acid region in the MC1R which will be discussed below.

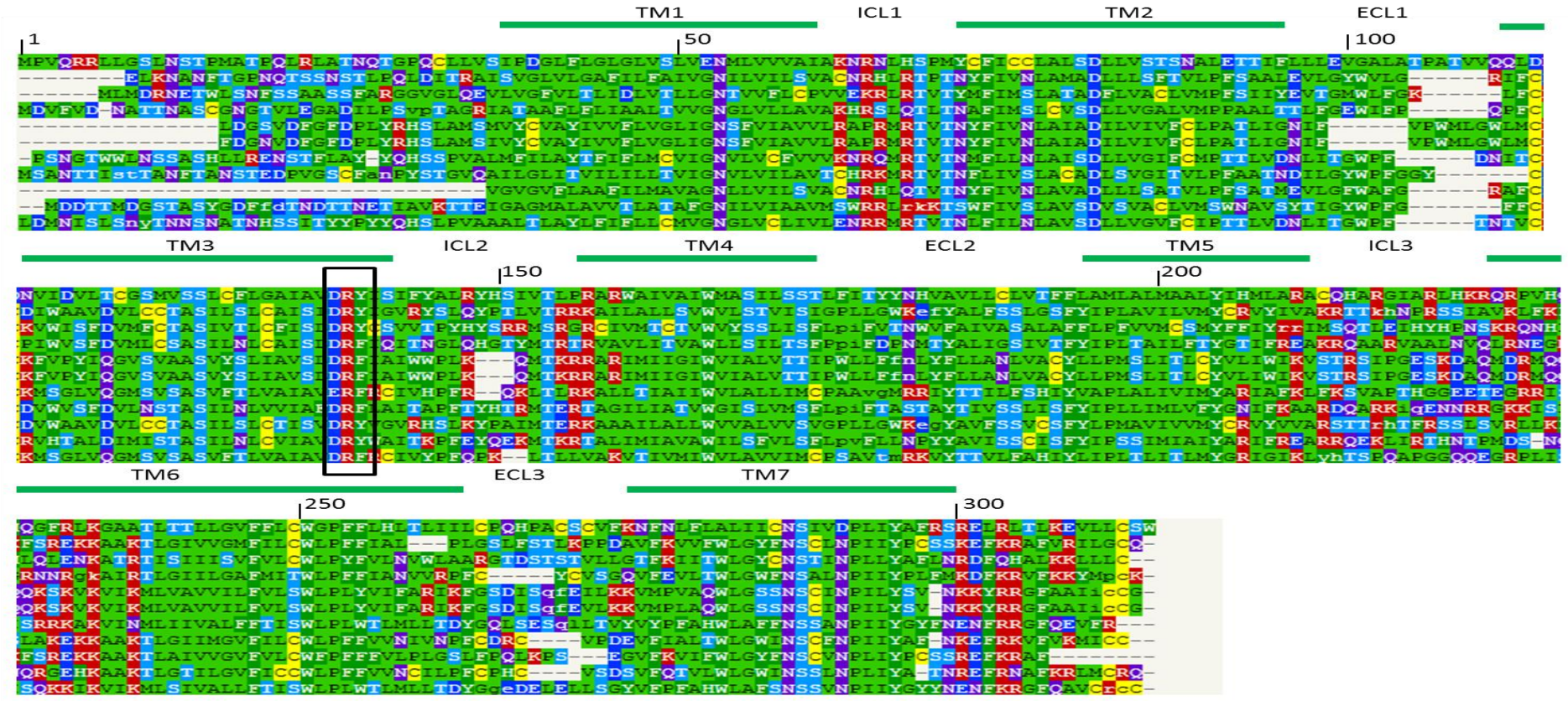


Figure 5.2. PSI-BLAST sequence alignment of the MC1R-wt obtained and adapted from Phyre2

Detailed view of the PSI-BLAST alignment with the MC1R-wt sequence on the top. The first ten sequences from the Phyre2 alignment are shown. TM = transmembrane domains, ICL = intracellular loops and ECL = extracellular loops. Colours: yellow = cysteine, red = lysine or arginine, purple = asparagine or glutamine, dark blue = glutamic acid or aspartic acid, light blue = threonine or serine. All other amino acids are green. The boxed area shows the conserved DRY motif. Numbers refer to amino acid numbers from M of the N terminus.

The next step in the Phyre2 analysis was the identification of the protein of known structure whose own PSI-BLAST most closely matches, in a statistical sense, the PSI-BLAST of the MC1Rs. Using this method, Phyre2 matched both the MC1R-wt and MC1RΔ24 to the adenosine A_{2a} receptor (A2AR) (PBD ID: 3EML, chain A). Although there is only 30% sequence identity between the MC1Rs and the A2AR, Phyre2 was able to model 85% of the MC1R sequence (residues 36-314) using the A2AR template. Phyre2 was unable to predict the structure of the N terminal part of the receptors (residues 1-35) and therefore this was not included in the pdb files obtained. Figure 5.3 shows a detailed prediction from Phyre2 of the secondary structure of the MC1R-wt and MC1RΔ24, based on the *known* secondary structure of the A2AR template. Although I-TASSER uses a different algorithm, it independently matched both the MC1R-wt and MC1RΔ24 to exactly the same A2AR template. These results are also consistent with the GPCR superfamily evolutionary tree shown earlier (figure 1.28), where the GPCR of known structure which is most closely related to the MC1R is indeed the A2AR.

Figure 5.3. Part A. Detailed sequence alignment obtained from Phyre2 showing the MC1R-wt sequence as query sequence and the human adenosine A2A receptor sequence as the template

MC1R-wt alignment showing the SNALETTI sequence, which is deleted from the MC1RΔ24, highlighted in a pale red box and the predicted extracellular loop one (ECL1) highlighted. Predicted secondary structure is indicated using green spirals for alpha helices, blue arrows for β strands and grey lines as unknown or disordered structure. Amino acids highlighted in grey are matched sequences. The dark red blocks show insertions relative to the template and the orange blocks show deletions relative to the template. The large orange bars denote a mismatch caused by the insertion of the T4 lysozyme added to the A2AR to form a chimera for optimal crystallisation. The amino acids boxed in red refer to catalytic sites of the T4 lysozyme and are of no relevance here.

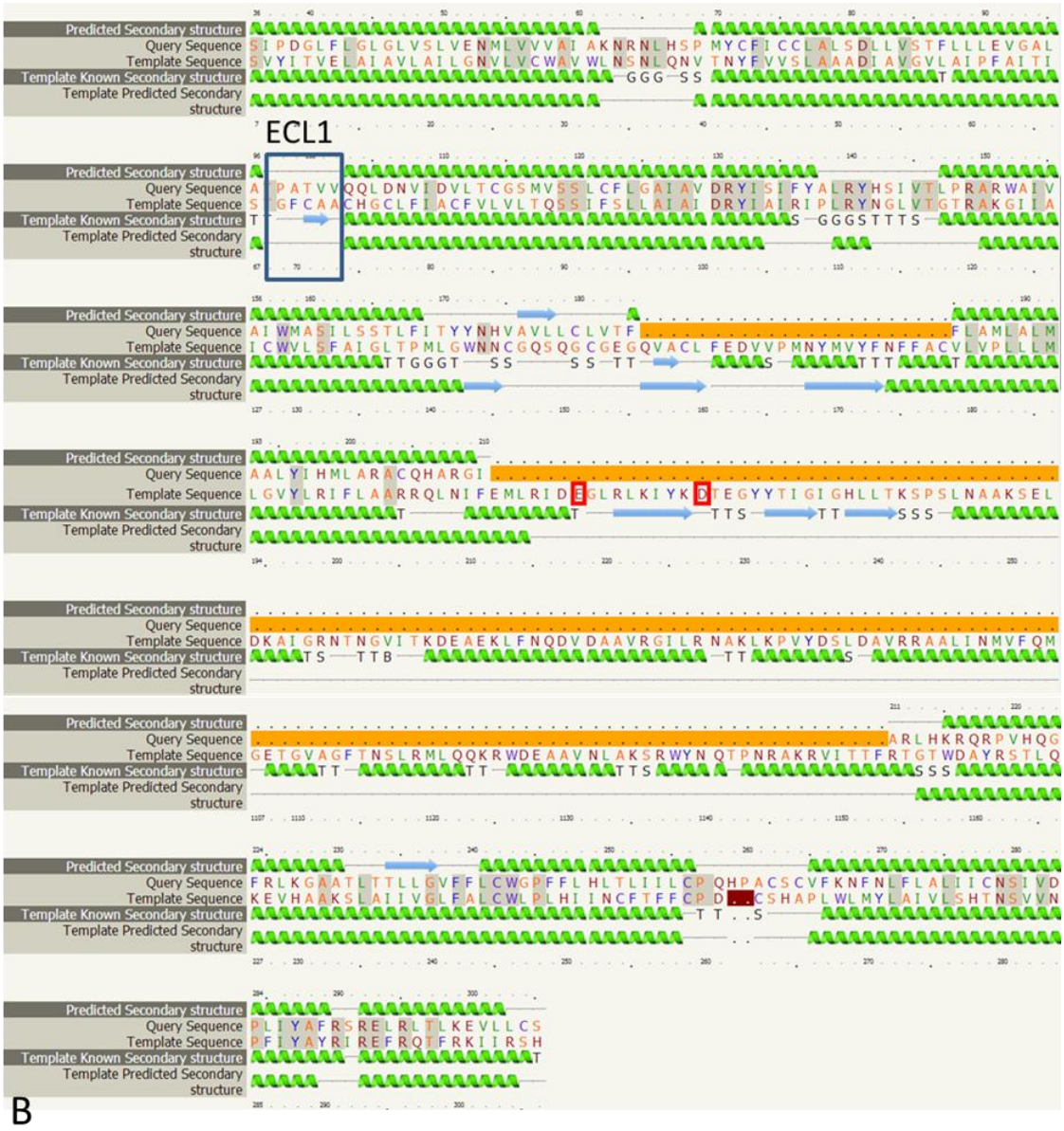


Figure 5.3. Part B. Detailed sequence alignment obtained from Phyre2 showing the MC1RΔ24 sequence as query sequence and the human adenosine A2A receptor sequence as the template

MC1RΔ24 alignment showing the predicted ECL1 highlighted. Predicted secondary structure is indicated using green spirals for alpha helices, blue arrows for β strands and grey lines as unknown or disordered structure. Amino acids highlighted in grey are matched sequences. The dark red blocks show insertions relative to the template and the orange blocks show deletions relative to the template. The large orange bars denote a mismatch caused by the insertion of the T4 lysozyme added to the A2AR to form a chimera for optimal crystallisation. The amino acids boxed in red refer to catalytic sites of the T4 lysozyme and are of no relevance here.

Figure 5.4 shows a sequence alignment from I-TASSER of the ten closest matches of the MC1R varaints to proteins of known structure, the top five all being the A2AR. The alignment illustrates some of the difficulties and limitations of creating accurate alignments. The area highlighted in red of figure 5.4 is shown in close up in figure 5.5. Although the second, third, fourth and fifth sequences are all identical, the alignments differ. For example, the second sequence has *two* gaps (98-99 and 105-110), the third sequence has *one* gap (100-107), while the fourth sequence has *two different* gaps (95-96 and 119-124). This is particularly relevant here as, these sequences are adjacent to the SNALETTI sequence deleted from the MC1RΔ24 (highlighted in blue in figure 5.5).

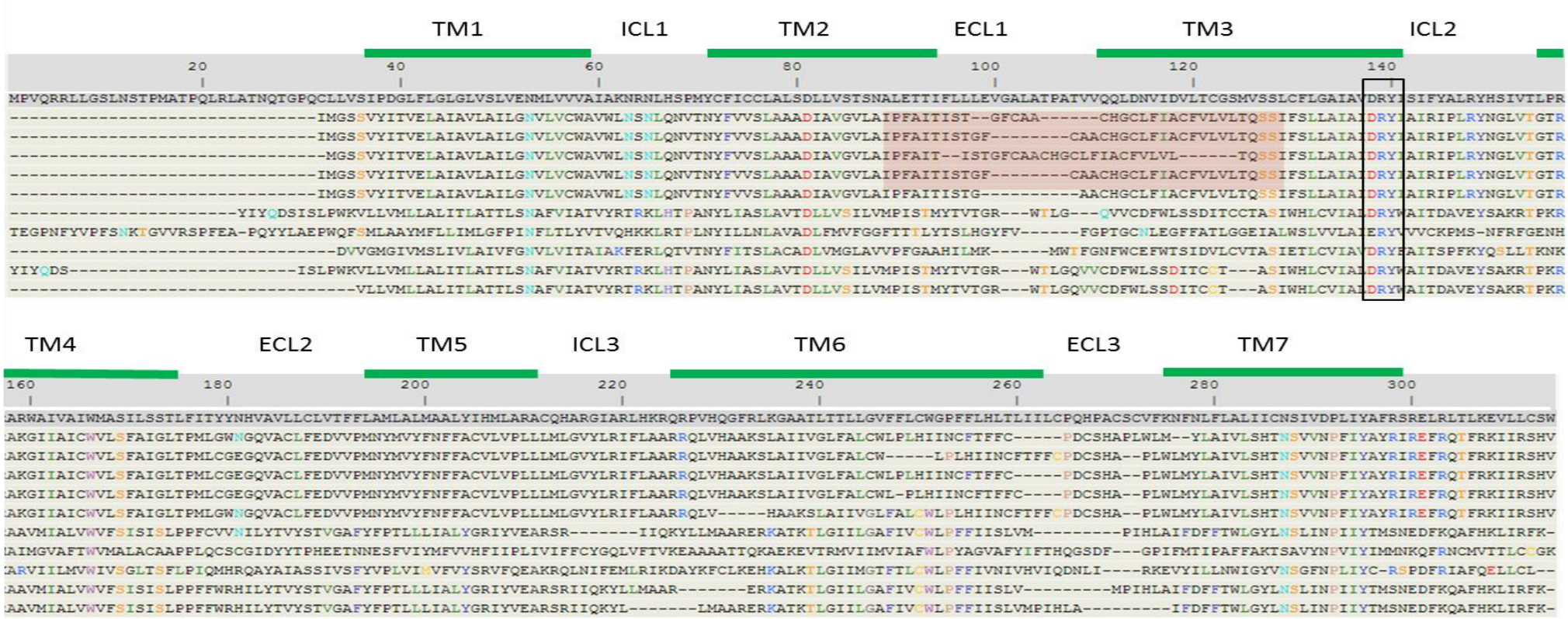


Figure 5.4. Sequence alignment of the MC1R-wt obtained and adapted from I-TASSER

The amino acid sequence at the top is the MC1R-wt and the sequences below show the top ten matches with the best match being the human adenosine A2A receptor (A2AR). Amino acids which match the MC1R are coloured and all others are black. The boxed area shows the conserved DRY motif. The area highlighted in red is discussed in figure 5.5.

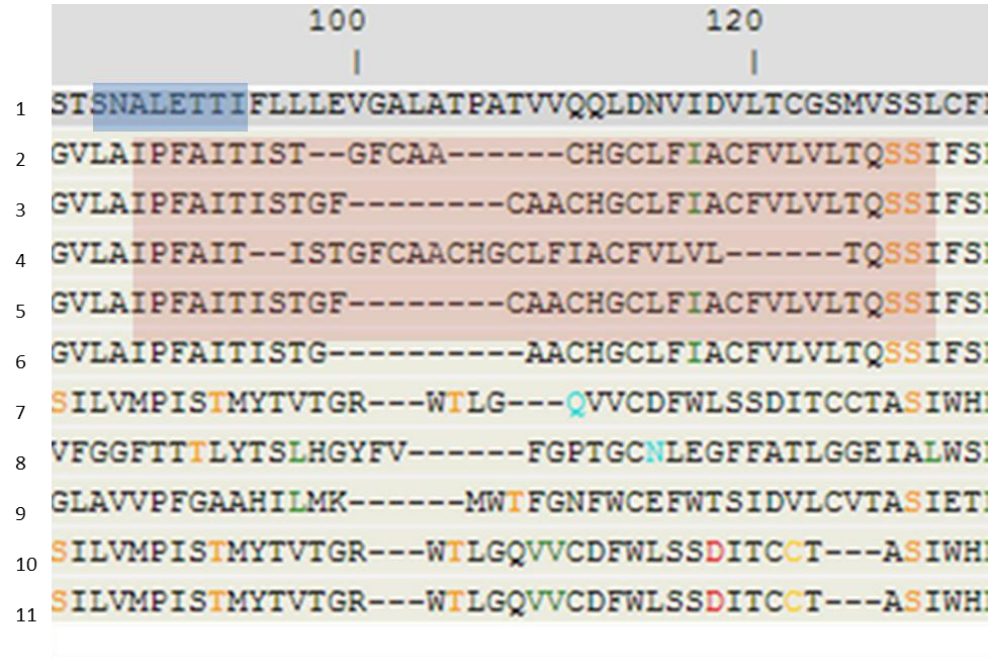


Figure 5.5. Detail of sequence alignment obtained and adapted from I-TASSER

The sequences highlighted in red are identical but have been aligned in such a way as to create different gaps. The arrow shows the position on the fourth sequence where a gap could have been created so as to match the third sequence. The SNALETTI sequence deleted in the MC1RΔ24 is highlighted in blue.

Both Phyre2 and I-TASSER created three dimensional computer models by threading the MC1R amino acids onto the A2AR template structure. Both Phyre2 and I-TASSER models predicted that the deletion of the SNALETTI (87-94) sequence from the MC1R-wt causes the FLLLEVGALA (95-104 MC1R-wt numbering) sequence of the extracellular loop 1 (ECL1) to be incorporated into the TM2 in the MC1RΔ24. This incorporation is predicted to result in a truncated ECL1 in the mutant, as shown in the three dimensional model of the MC1R variants in figure 5.6. Although Phyre2 and I-TASSER models both used the same template, the pdb files differ in fine details, as is illustrated later when investigating the residues of the binding site. However, for consistency and conciseness, only the Phyre2 models are used in subsequent analyses.

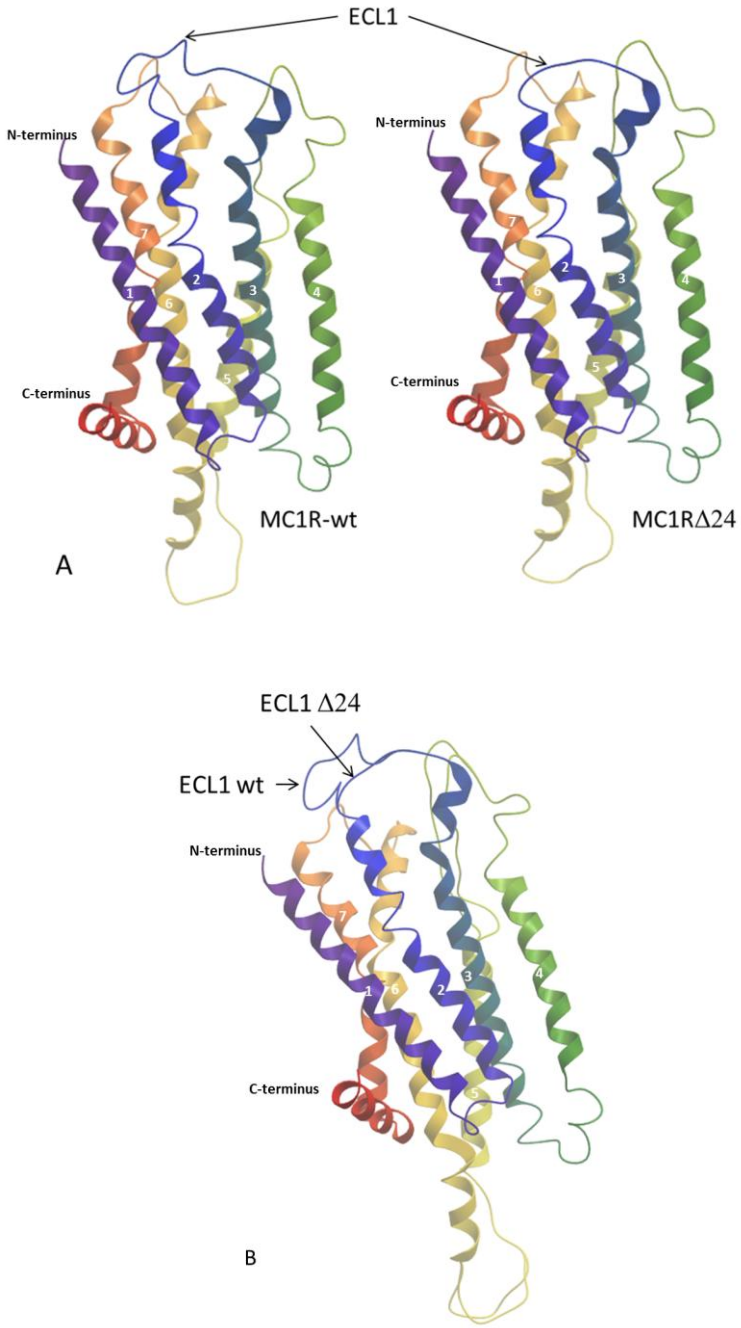


Figure 5.6. Computer models of the MC1R-wt and MC1RΔ24 side view predicted by Phyre2

The model is coloured using the rainbow spectrum from blue at the N- terminus to red at the C-terminus. Numbers one to seven refer to the transmembrane helices. A) Shows MC1R-wt and MC1RΔ24 side by side with the extracellular loops 1 (ECL1) highlighted and B) shows the two receptors superimposed on each other. Models cover 85% of the structure of the MC1R, the extracellular N-terminus region not being predicted in this case. Pdb models visualised with Molsoft.

Figure 5.7 shows that the ECL1 of the MC1R-wt is predicted to have a section of β strand which is missing in the MC1RΔ24.

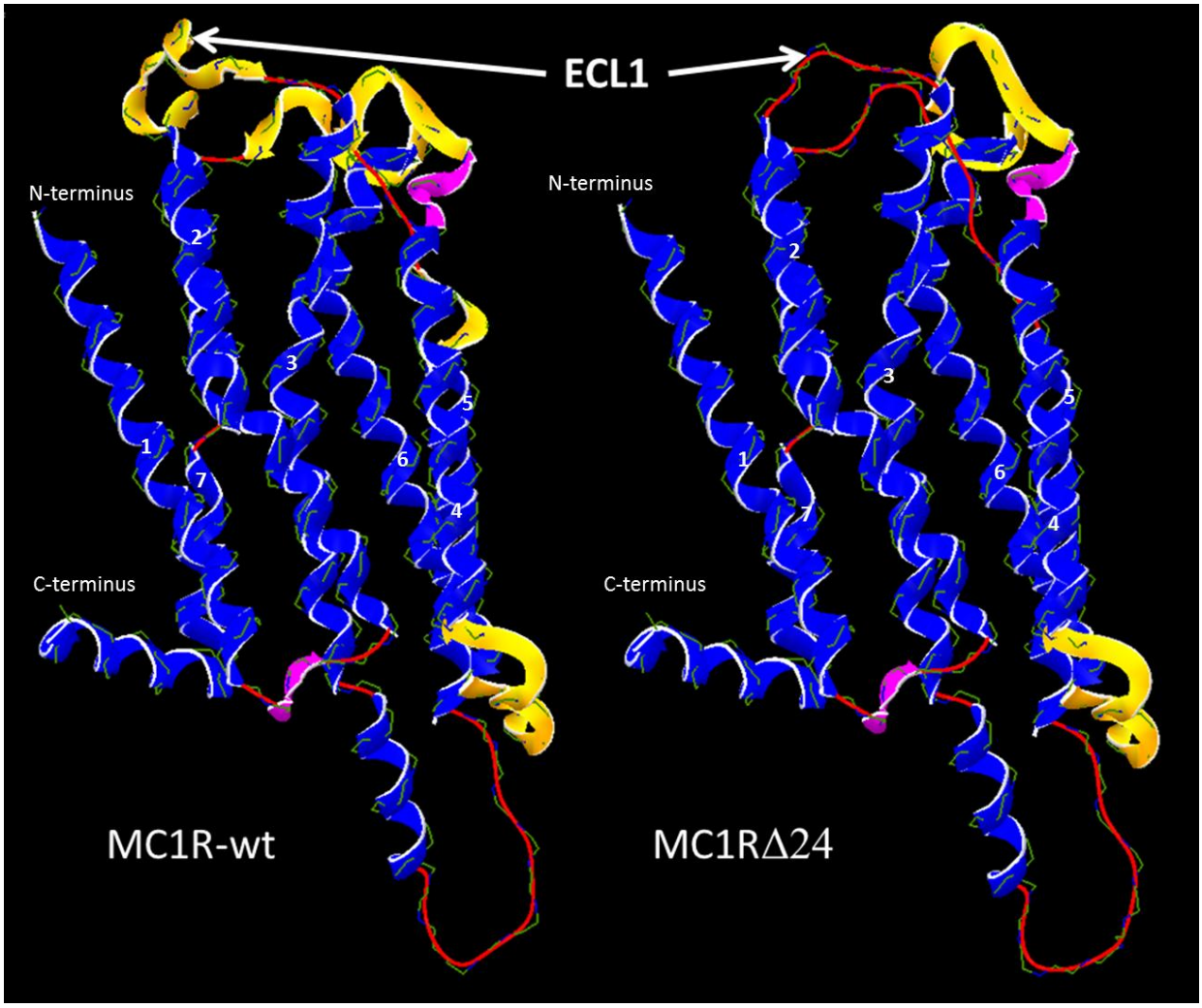


Figure 5.7. Computer models of the MC1R-wt and MC1RΔ24 side view predicted by Phyre2

Computer models predicted by Phyre2 visualised with Hex 8.0 (Macindoe et al., 2010). Extracellular loop 1 (ECL1) is indicated showing the β strand, represented in yellow, present in the wildtype and absent in the mutant receptor. Blue represents alpha helices and red lines are non-repetitive loops. Numbers refer to the transmembrane helices one to seven. Models cover 85% of the structure of the MC1R, the extracellular N-terminus region not being predicted in this case.

The model visualised with Hex 8.0 was used to generate a schematic, two dimensional representations of the receptor variants, as shown in figure 5.8, with the top side representing the extracellular side of the receptor. The predicted disulphide bridges and the highly conserved DRY motif are shown.

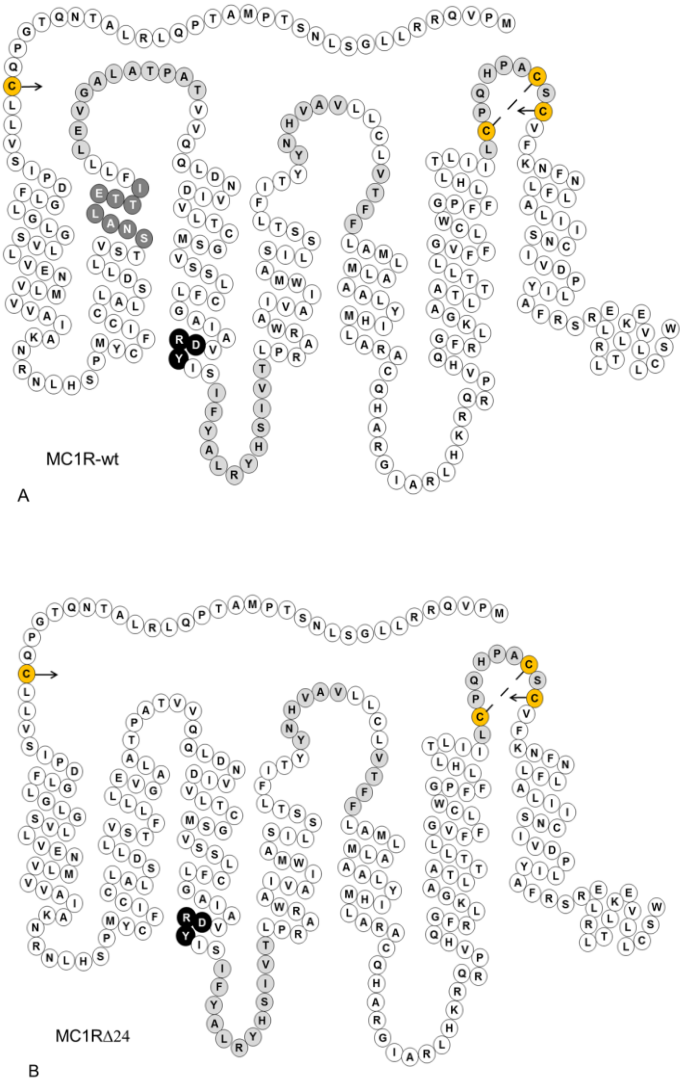


Figure 5.8. Schematic representation of the MC1R-wt and MC1RΔ24 based on Phyre2 and Hex 8.0 predictions

Amino acid residues that form the β strands are shown as light grey circles. A) Predicted structure of the MC1R-wt showing the eight amino acids (SNALETTI) deleted in the MC1RΔ24 as dark grey circles with white lettering. The highly conserved DRY motif of transmembrane helix three is shown as black circles with white lettering. B) Predicted structure of the MC1RΔ24 showing a shortened extracellular loop 1 (ECL1) with the FLLLEVGALA sequence having been relocated to the top of transmembrane helix two. Cysteines involved in disulphide bridges are shown in orange with arrows or broken lines to show which cysteines are connected in each case.

5.2.2 Melanocortin-1 receptor intramolecular interactions

Computer models predicted that there were very few differences in the structures of the seven transmembrane helices between the wildtype and mutant receptors. There was one hydrogen bond difference detected where the wildtype receptor has a hydrogen bond between N88 (of SNALETTI) of TM2 and S124 of TM3. This hydrogen bond is missing in the mutant receptor, as shown in figure 5.9. There were other hydrogen bonds and other structural differences in the ECLs of the receptors but these are not considered reliable here due to the limitations of homology modelling for extracellular loops.

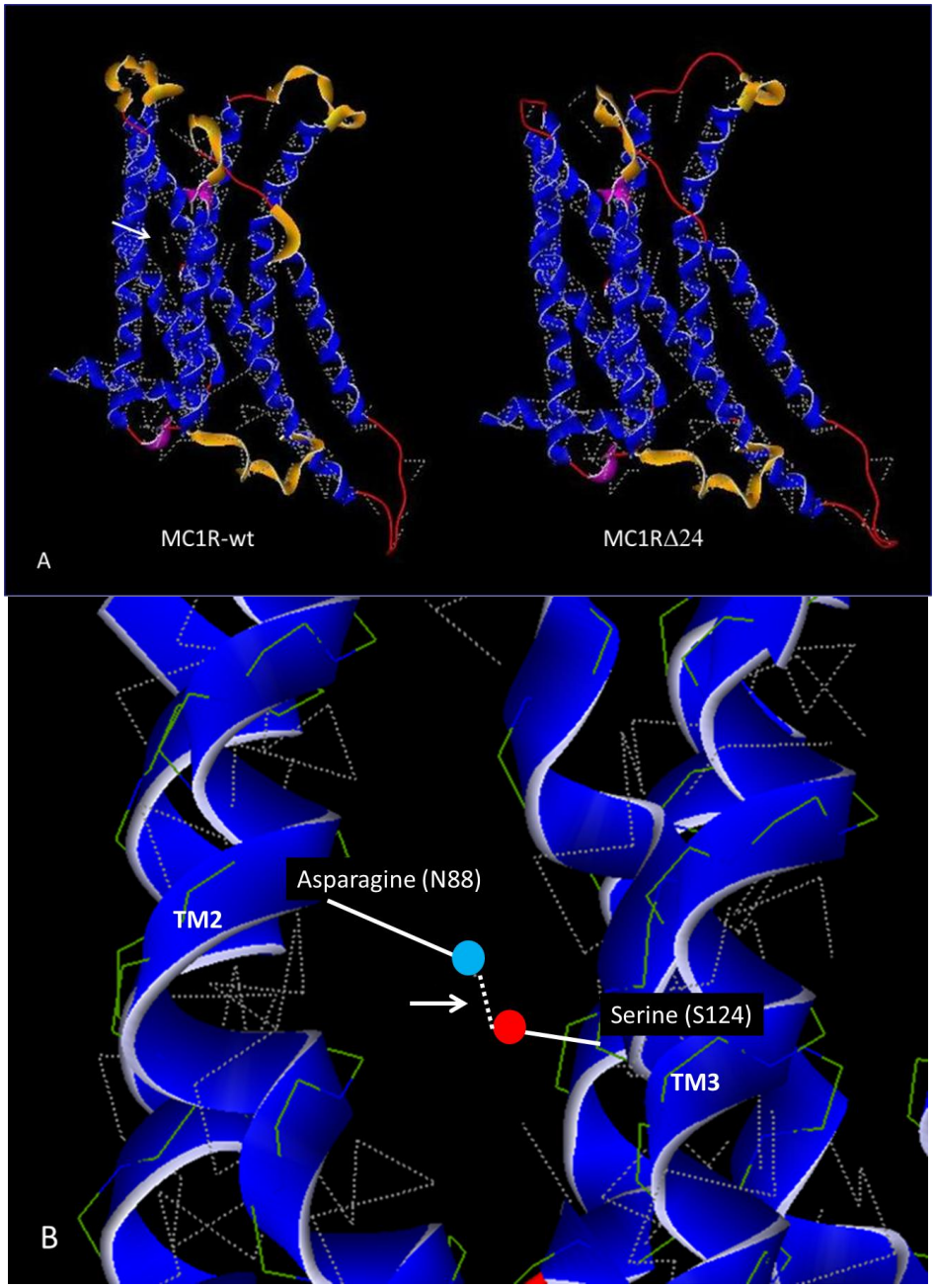


Figure 5.9. Computer model of the MC1R-wt and MC1RΔ24 visualised with Hex 8.0

Hydrogen bonds are shown as dotted white lines. A) MC1R-wt and MC1RΔ24 with the hydrogen bond between N88 and S124 indicated with an arrow in the wildtype and absent in the mutant. B) Close up of the hydrogen bond in the MC1R-wt. The side chains of N88 and S124 are represented as white lines with nitrogen of the N88 as a blue circle and oxygen of the S124 as a red circle.

The pdb files obtained from Phyre2 were used to further analyse the structural similarities and differences between the wildtype and mutant receptor using the computer programme Socket to investigate the “knobs into holes” interactions between helices as described in the introduction (Walshaw and Woolfson, 2001). “Knobs into holes” interactions occur where hydrophobic amino acids from one helix become interlocked into hydrophobic holes created by amino acids on an adjacent helix. Results from this programme showed that most interactions are towards the intracellular side of the molecules and that both wildtype and mutant receptors have identical interactions. Figure 5.10 shows the residues involved in the interactions.

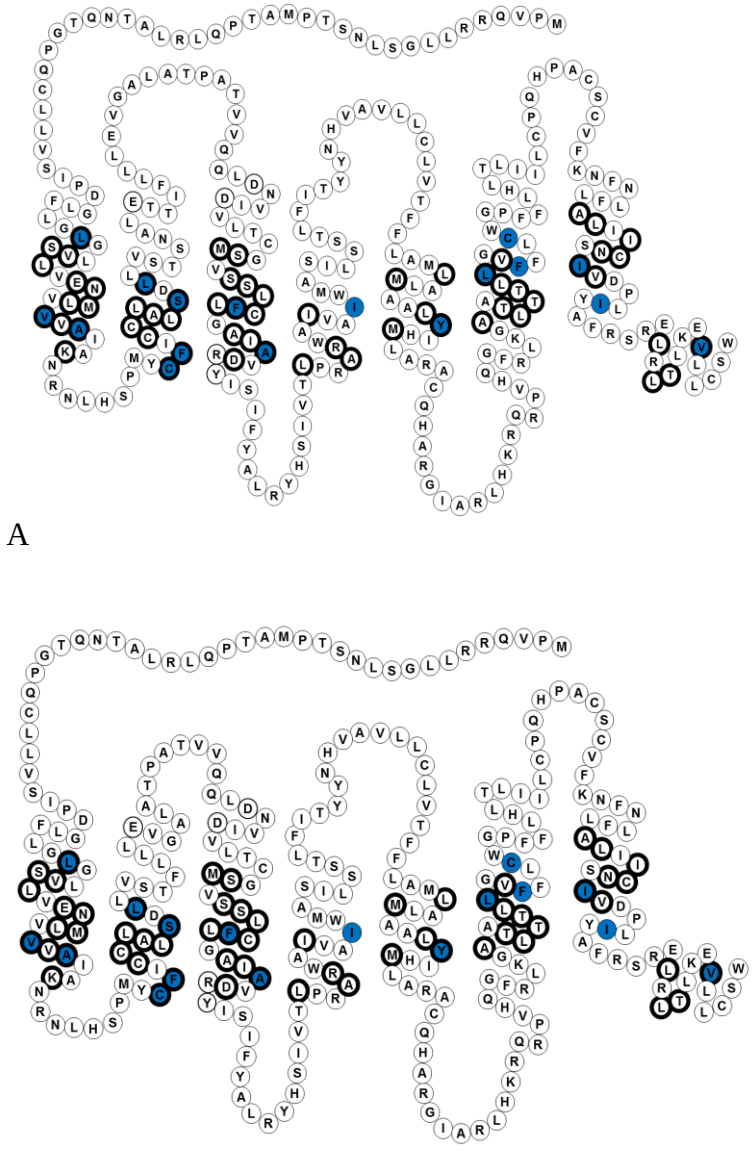


Figure 5.10. Schematic representations of A) MC1R-wt and B) MC1RΔ24 showing residues predicted to be involved in "knobs into holes" interactions

"Knobs into holes" interactions occur where hydrophobic amino acids from one helix become interlocked into hydrophobic holes created by amino acids on an adjacent helix. The blue residues represent "knobs" and the residues with bold outlines represent "holes".

Figure 5.11 shows which helices are predicted to interact in a "knobs into holes" fashion.

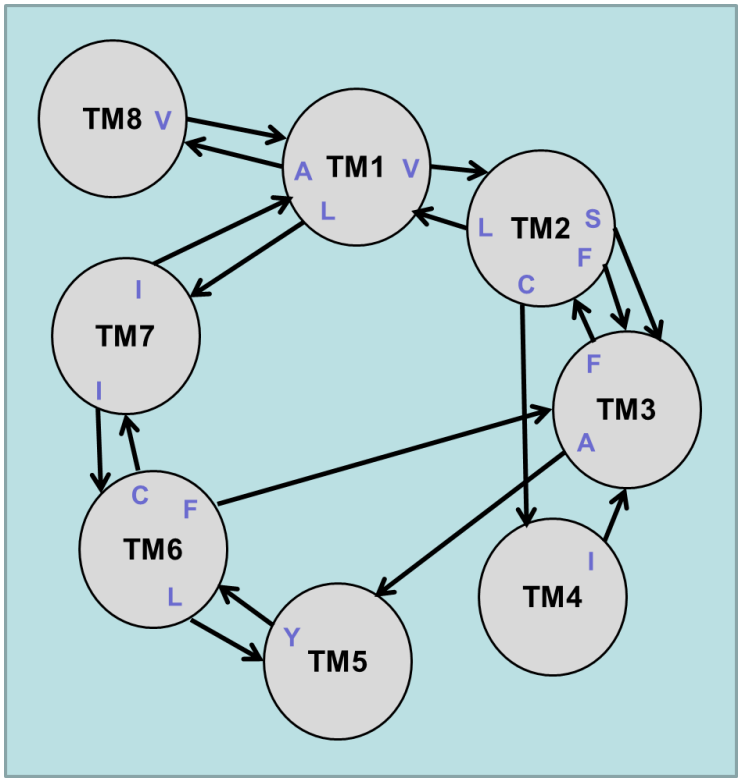


Figure 5.11. Transmembrane helices (TM) of the melanocortin-1 receptor showing the predicted "knobs into holes" interactions

Representation of the MC1R-wt viewed from above, with circles as transmembrane helices. "Knobs into holes" interactions occur where hydrophobic amino acids from one helix become interlocked into hydrophobic holes created by amino acids on an adjacent helix. Socket predictions of interactions between helices are represented as arrows. Letters refer to residues which act as "knobs".

Details of the specific interactions which are predicted by Socket are shown in appendix 4.

5.2.3 Melanocortin-1 receptor binding pockets

The MC1R has two binding pockets, one on the extracellular side of the receptor for ligand binding and the other on the intracellular side of the receptor for G-protein binding. Figure

5.12 shows the two binding pockets of the MC1R-wt and MC1RΔ24 receptors highlighted in green as predicted by *CLC Discovery* (Thomsen and Christensen, 2006). The extracellular binding sites of the MC1R-wt and MC1RΔ24 are predicted by *CLC Discovery* to be 313 $Å^3$ and 256 $Å^3$ respectively. The intracellular binding sites for both variants are 156 $Å^3$ (Thomsen and Christensen, 2006). These binding site volumes are calculated by the "rolling ball" algorithm, using a probe of 1.4 Å, which approximates the radius of a water molecule, to represent a solvent (Shrake and Rupley, 1973). This probe is "rolled" over the surface of the molecule allowing the volume of the pocket to be calculated.

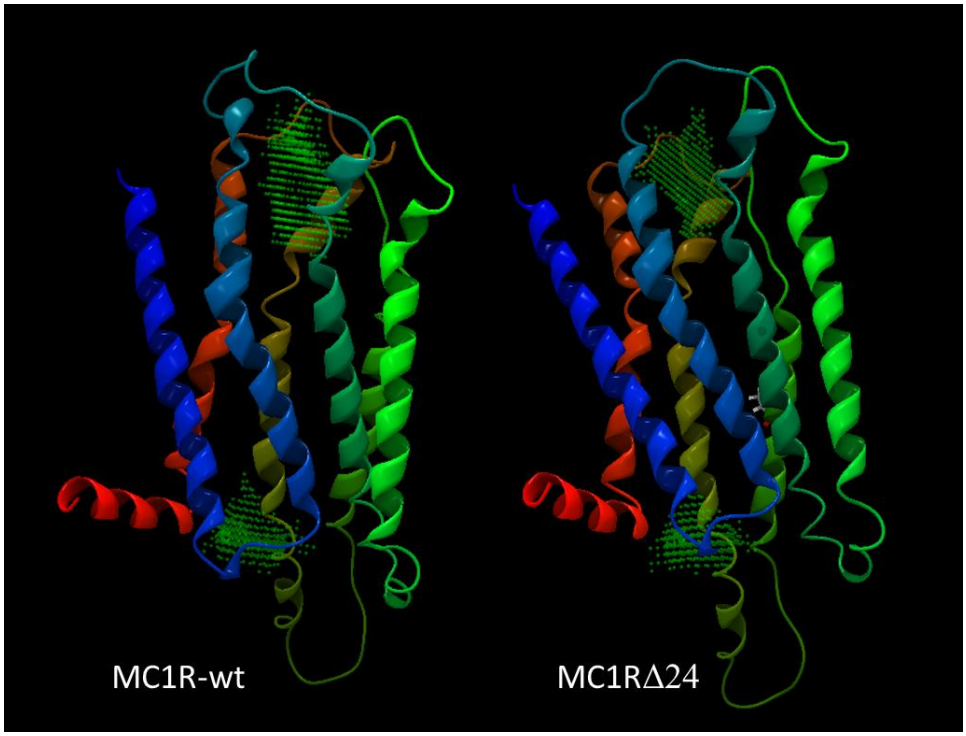


Figure 5.12. Computer models of the MC1R-wt and MC1RΔ24 from *CLC Discovery* showing the predicted binding pockets in green dots

Figure 5.13 shows the extracellular view of the binding pockets shown as A) a ribbon structure and B) the molecular surface. The molecular surface uses the same "rolling ball" algorithm, described above, to create a representation of the surface of a molecule, as seen from the point of view of a water molecule. Figure 5.13 B highlights how ECL1 from the MC1R-wt partially covers the binding site. This may impact on the ability of a ligand to enter the binding pocket. In contrast, ECL1 of MC1RΔ24 appears to leave the binding site more exposed and possibly more accessible to ligands.

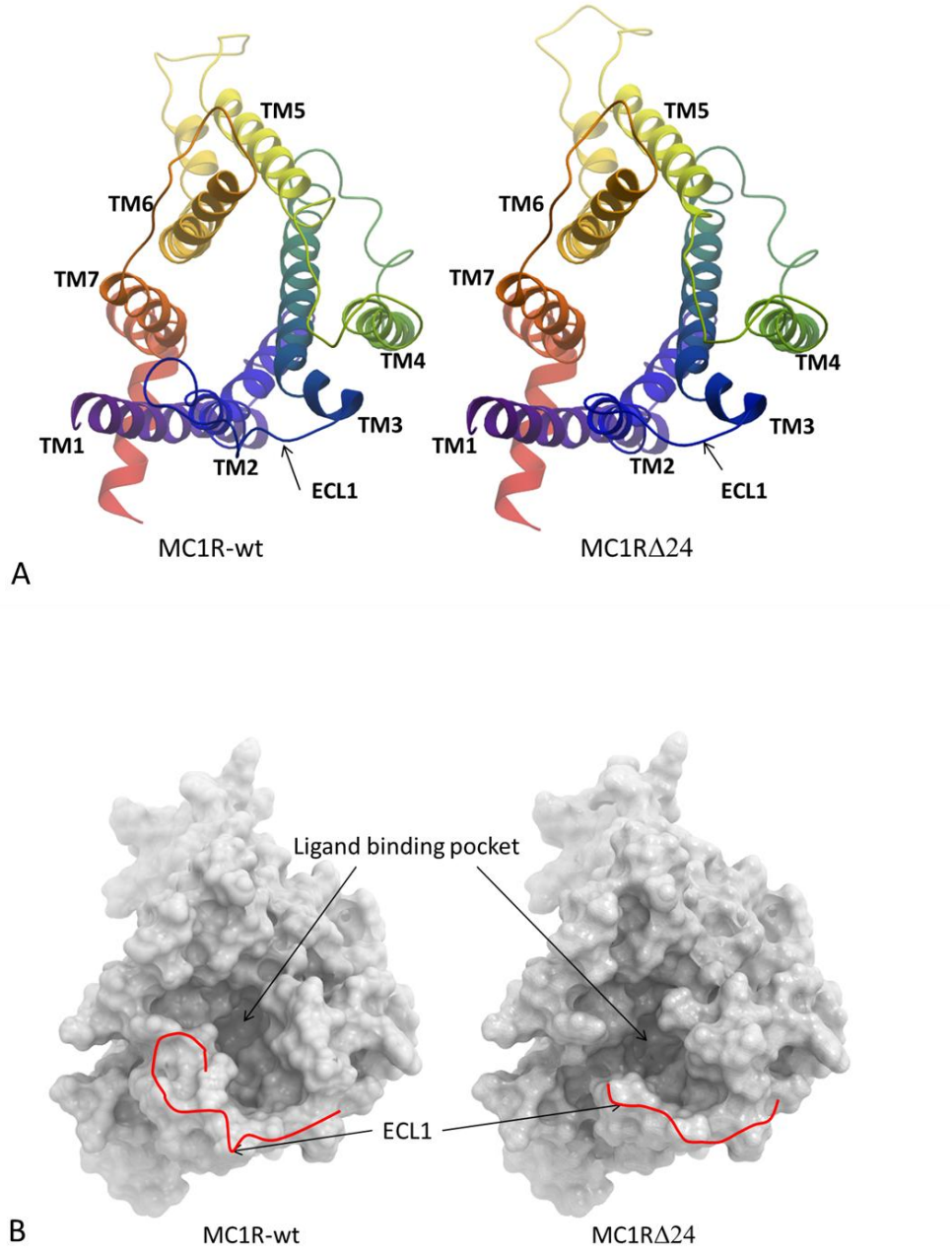


Figure 5.13. Computer predictions of the extracellular views of binding pockets of the MC1R-wt and MC1RΔ24

A) Ribbon diagram representation of the MC1R-wt and MC1RΔ24. B) The same orientation as the ribbon diagram, represented as a molecular surface. The red lines indicate the positions of extracellular loops in each case. TM= transmembrane helices, ECL1= extracellular loop one.

The intracellular view of the G-protein binding site is shown in figure 5.14. The wildtype and mutant binding sites are essentially identical according to these models with only slight variation in the positioning of the intracellular loops.

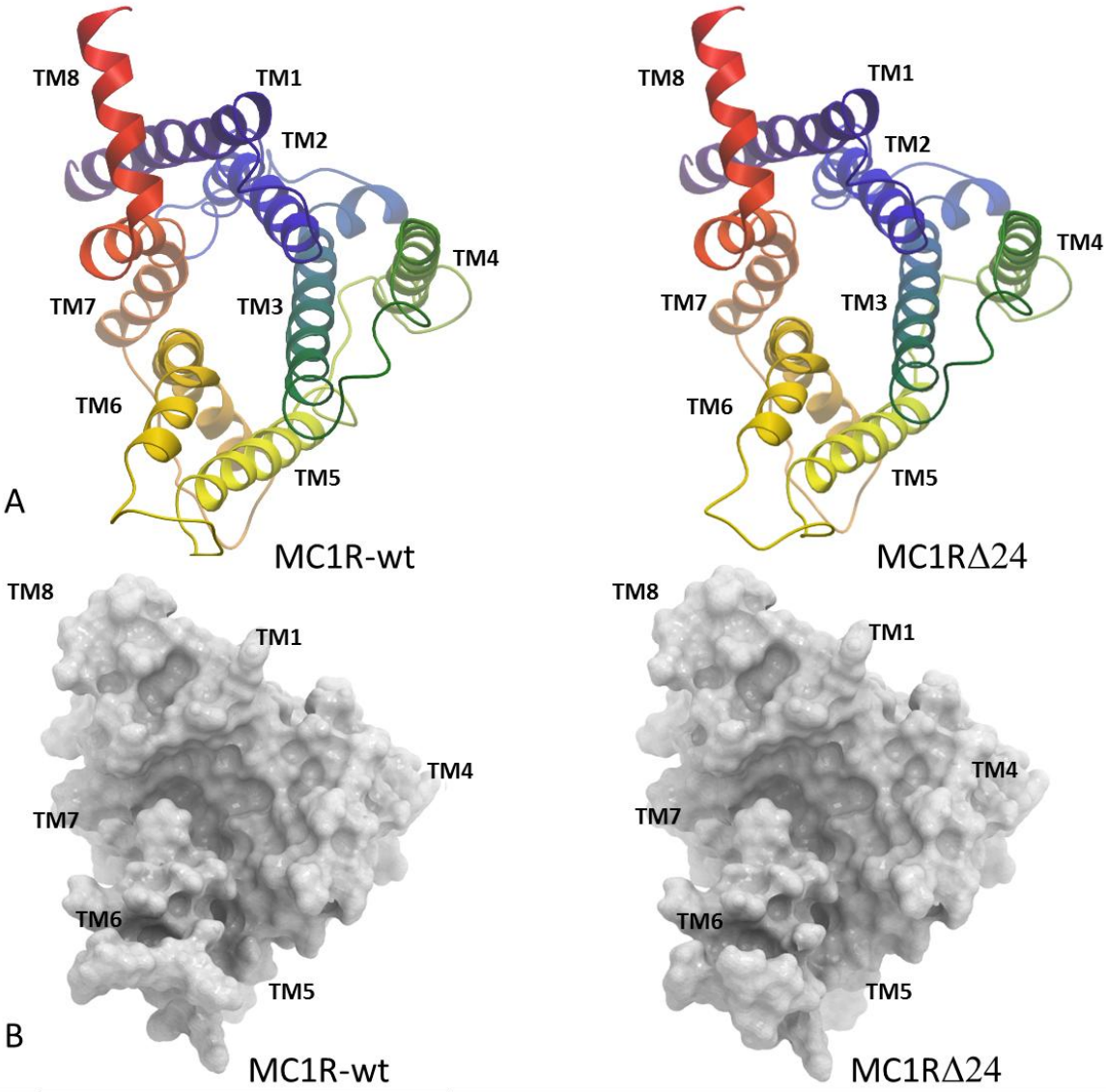


Figure 5.14. Computer models showing the intracellular view of the G-protein binding sites of MC1R-wt and MC1RΔ24

A) Ribbon diagram representation of the MC1R-wt and MC1RΔ24. B) The same orientation as the ribbon diagram, represented as a molecular surface.

Figure 5.15 A and B show the extracellular view of each receptor with electrostatic surfaces. The red area shows a region of negative charge in the binding pocket of both receptors. Figure 5.15 C and D show the same view as a ribbon diagram with the negatively charged residues E91, D114 and D118 highlighted (MC1R-wt numbering). These residues are thought to be involved in ligand binding in a number of species, where they form an acidic pocket (Lu, Vage and Cone, 1998; Yang, 2011). The specific residues of the binding pocket of the MC1R-wt and MC1RΔ24 are considered next.

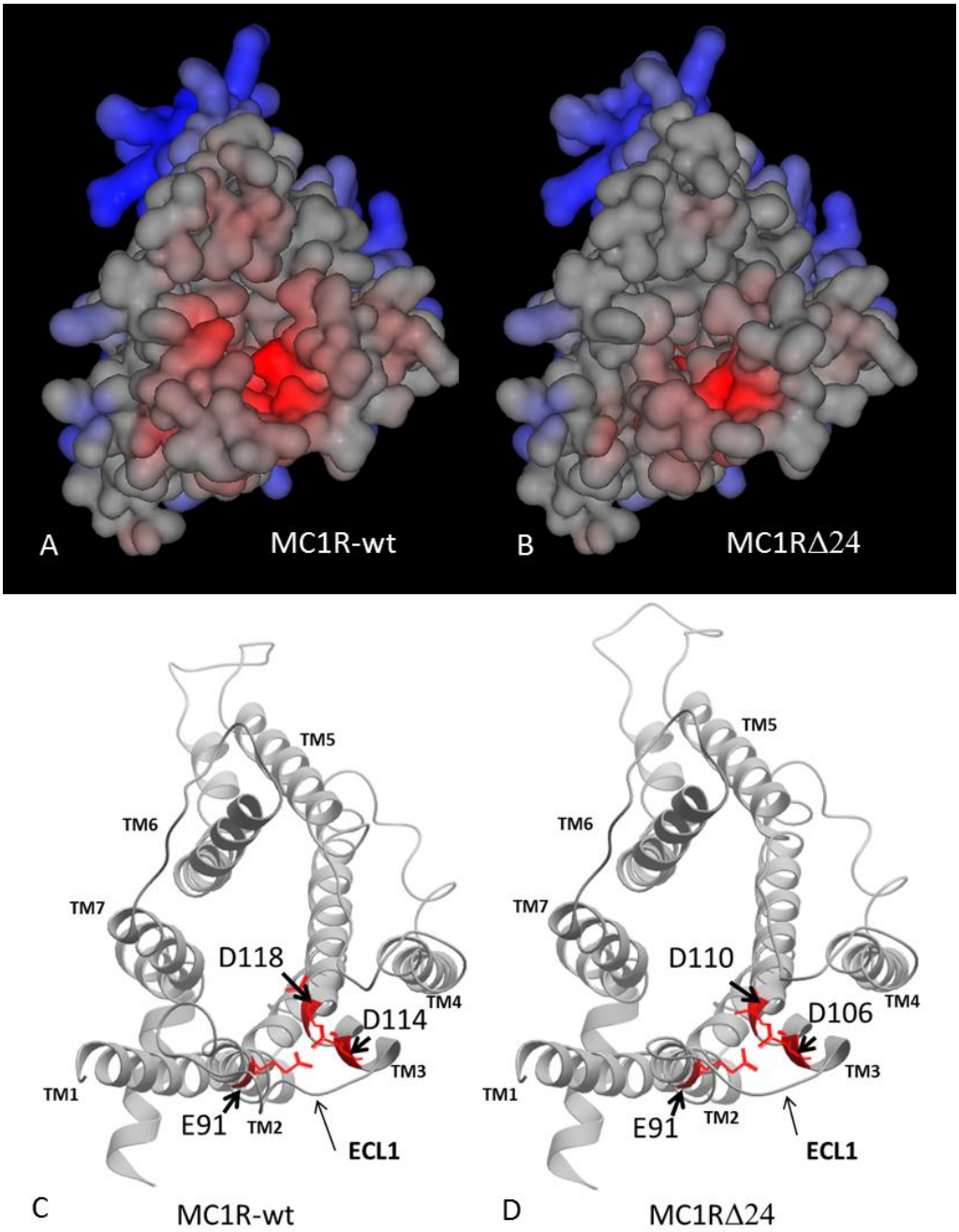


Figure 5.15. Extracellular view of the MC1R-wt and MC1RΔ24 showing the binding pockets

A) and B) show solid molecular surfaces, with blue representing positive charges and red representing negative charges and with intermediate charges in grey. C) and D) show ribbon diagrams of the same view as A and B with the E91, D114 and D118 (MC1R-wt numbering) residues highlighted as stick models (MC1R-wt numbering). A and B are visualised with Hex 8.0 (Macindoe et al., 2010) *and C and D are visualised with Molsoft.*

5.2.4 Residues found in the binding site

In order to gain an understanding of the MC1R extracellular binding pocket and the residues that may be involved in ligand interaction, three approaches were taken. Firstly, data were obtained from the literature on experiments with natural mutants and mutagenesis investigations involving the MC1R and human MC4R. Secondly, data were obtained from the crystallographic analysis of the A2AR, this being the template structure used by Phyre2 and I-TASSER to build the MC1R models in the present study. Thirdly, pdb files of the MC1R-wt and MC1RΔ24 were analysed with the computer programmes Fpocket and Pepsite to predict which residues were in the binding pocket.

5.2.4.1 Experimental data from the MC1R and MC4R

There are a number of natural and artificial MCR mutants that have been pharmacologically characterised. Specific residues were included in the data if they were found to be involved in receptor-ligand interactions and were also found in the vicinity of the receptor's binding pocket. These interactions could involve changes to activity and/or binding affinity. Naturally occurring MC1R residues thought to be involved include E92 and L98 (mouse) (mouse numbering) (Robbins et al., 1993; Benned-Jensen, Mokrosinski and Rosenkilde, 2011), C123 (fox) (mouse numbering) (Våge et al., 1997a) and D84, V92, L93 (human) (human numbering) (Más et al., 2002; Ringholm et al., 2004; Beaumont et al., 2005). Mutagenesis data on the human MC1R have shown that E94, D117, D121, T124 F175, Y183, D184, H185, F195, F196, F257, F258, F277, F280 and L284 (human numbering) are involved (Yang et al., 1997; Chai et al., 2005). Mutagenesis data on the mouse MC1R have confirmed that E92, D119 and C123 (mouse numbering) are involved (Lu, Vage and Cone, 1998).

Mutagenesis data on the human MC4R have shown that E91, D114. D118, T121, L133, S172, W251, F254, F257, L258 and F277 (human numbering) are involved (Yang et al., 2002; Chai et al., 2005). Mutagenesis data on the mouse MC4R have shown that E92, D114, D118, F176, Y179, F253, F254 and F259 (mouse numbering) are involved (Haskell-Luevano et al., 2001). All data are presented in appendix 5.

5.2.4.2 Experimental data from Adenosine 2A receptor

The A2AR receptor has been crystallised, bound to the synthetic agonist NECA and to caffeine (PDB ID: 2ydv and 3rfm respectively). Analyses of these structures suggest that when the A2AR receptor binds to the NECA, the following residues are involved in binding: V121, L122, T125, T188, C189, L197, L201, W251, L254, H255, N258, M277, I281, S284 and H285 (MC1R-wt numbering, A2AR residues) (Lebon et al., 2011). When the A2AR binds to caffeine, the following residues are thought to be involved in binding: T188, L197, L254, N258, M277 and I281 (MC1R-wt numbering, A2AR residues) (Doré et al., 2011). Figure 5.16 A,B,C,D, E and F shows the MC1R-wt with the ligand NECA superimposed in the predicted ligand binding pocket, as predicted by I-TASSER. Figure 5.16 B and C show the predicted binding pocket highlighted in pale green around the NECA ligand and the side chains which are predicted, from experimental evidence, to be interacting with the ligand (Lebon et al., 2011). Figure 5.16 G, H and I show the MC1R-wt with the ligand caffeine. The binding pockets are highlighted in pale green for each showing that the same receptor can have different binding pockets and bind diverse ligands. All data are included in appendix 5.

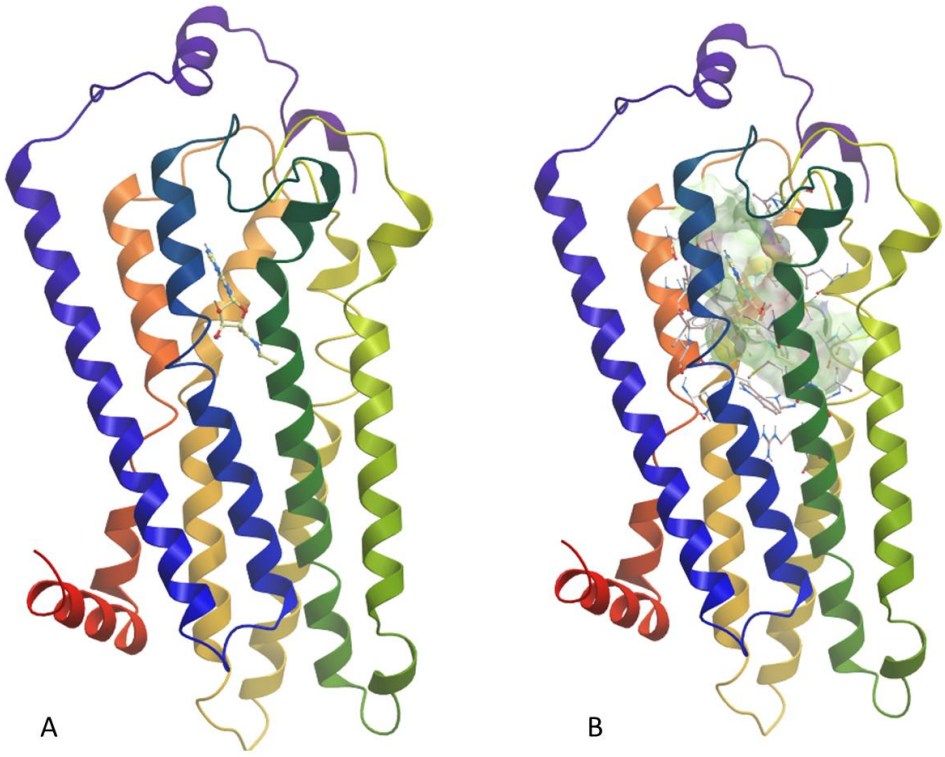


Figure 5.16. Part A and B. Computer models of MC1R-wt predicted by I-TASSER

A) MC1R-wt showing the position of the ligand NECA bound to the A2AR (PDB ID: 2ydy). B) The MC1R-wt showing the binding pocket for the NECA bound to A2AR.

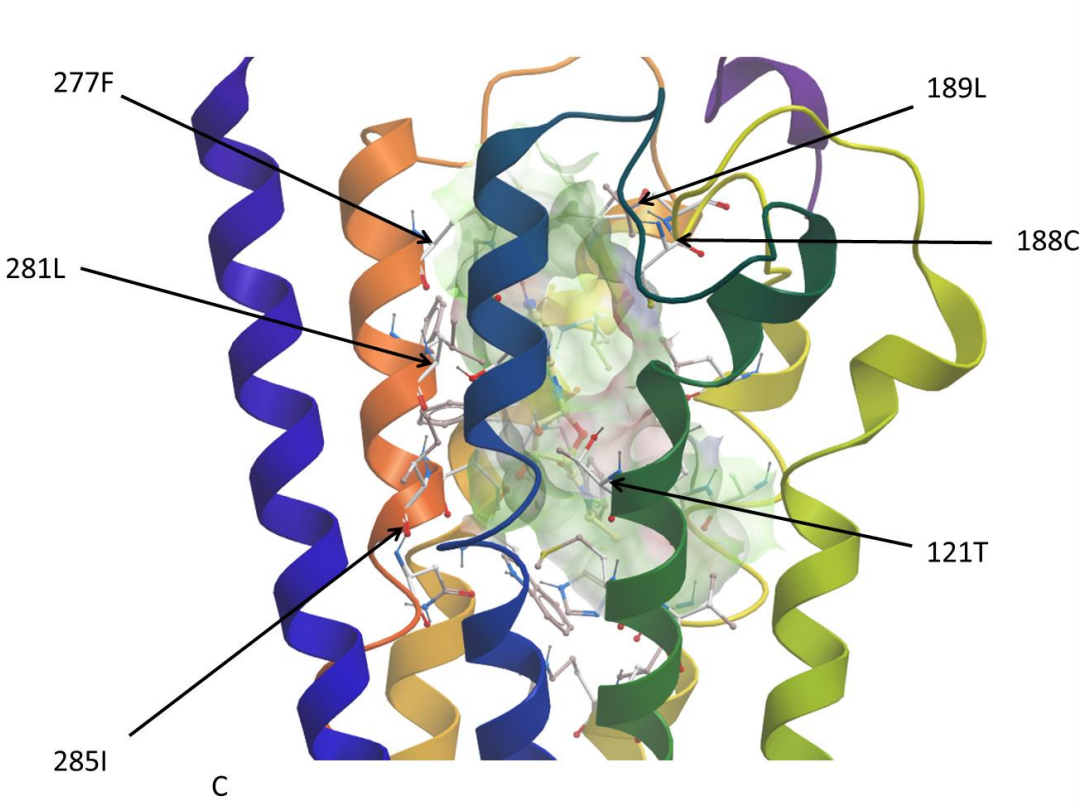


Figure 5.16. Part C. Computer models of MC1R-wt predicted by I-TASSER

C) MC1R-wt showing the binding pocket with side chains labelled.

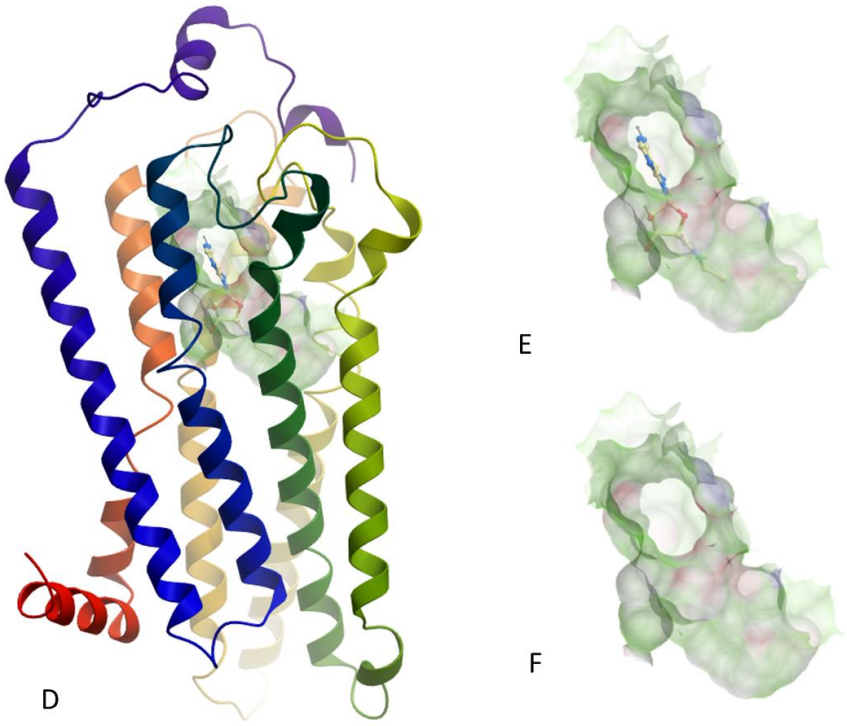


Figure 5.16. Part D, E and F. Computer models of MC1R-wt predicted by I-TASSER

D) MC1R-wt showing the position of NECA bound to the A2AR E) shows the binding pocket and NECA and F) shows the empty binding pocket.

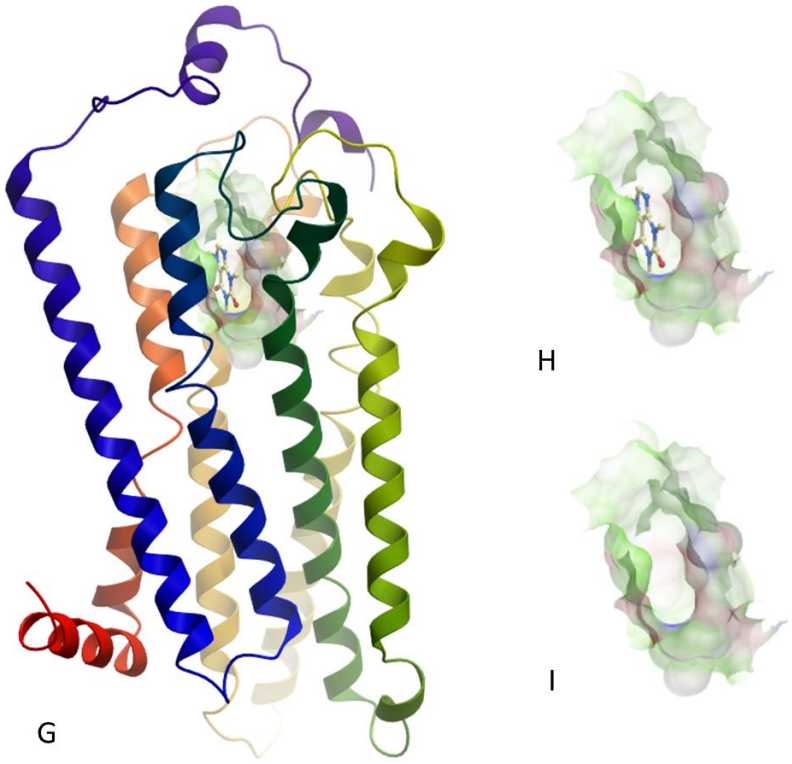


Figure 5.16. Part G, H and I. Computer models of MC1R-wt predicted by I-TASSER

G) MC1R-wt showing the position of caffeine bound to the A2AR (PDB: 3rfm), H) shows the binding pocket and caffeine and I) shows the empty binding pocket.

5.2.4.3 Data from the computer programmes Fpocket and Pepsite

The computer programme Fpocket used the MC1R pdb files to predict the position of the binding site in the receptor and which residues would be in the site. The programme found a range of residues for all the models provided, and results are shown in figure 5.17 and presented in appendix 5. Figure 5.17 shows that there is little consensus between the Phyre2 and I-TASSER predictions, highlighting the limitations of homology modelling. This finding suggests that computer predictions are less reliable than experimental data at predicting which residues may be in the binding site.

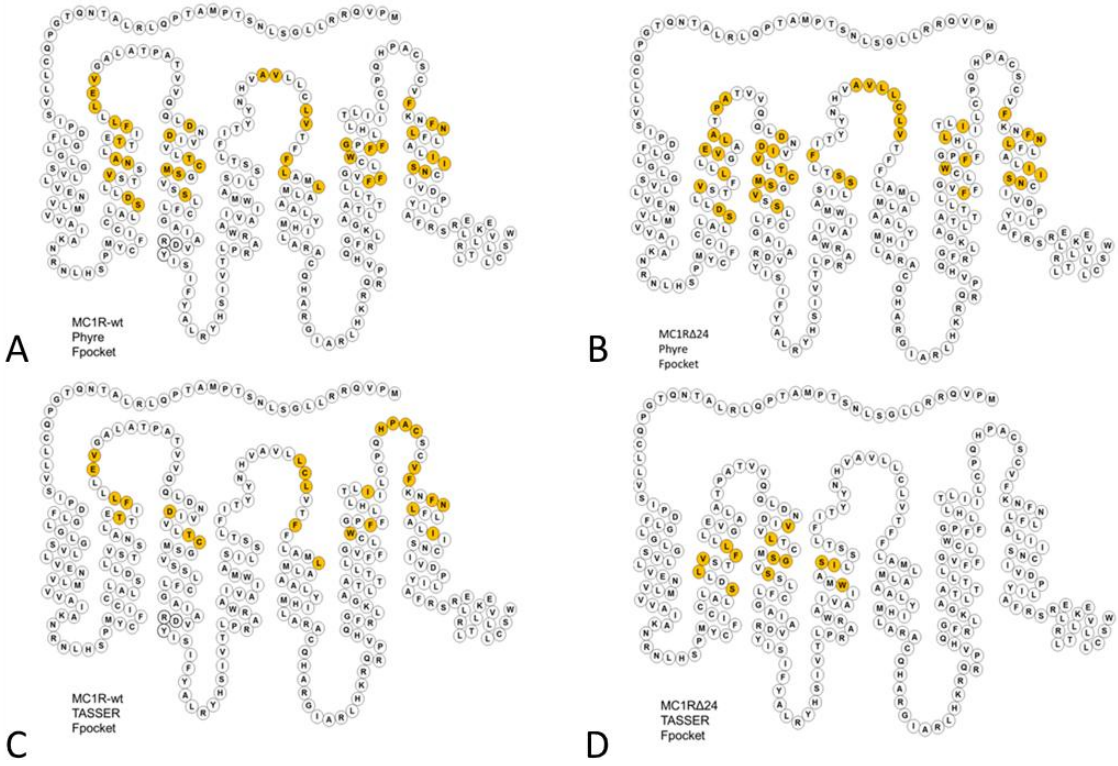


Figure 5.17. Schematic representations of the MC1R-wt and MC1RΔ24 showing predicted amino acids in the binding pocket highlighted in orange

Residues predicted to be in the receptor's binding site using the A) Phyre2 MC1R-wt model analysed with Fpocket B) Phyre2 MC1RΔ24 model analysed with Fpocket C) I-TASSER MC1R-wt model analysed with Fpocket D) I-TASSER MC1RΔ24 model analysed with Fpocket.

The web server Pepsite was used to give a prediction of the residues involved in specific binding with the HFRW motif, which is the essential core sequence of the α-MSH peptide. The process was repeated for the conserved RFF sequence of ASIP. The data obtained from Pepsite is particularly useful as it is receptor and ligand specific. The predicted interactions obtained from Pepsite are shown in figure 5.18 and the data are included in appendix 5.

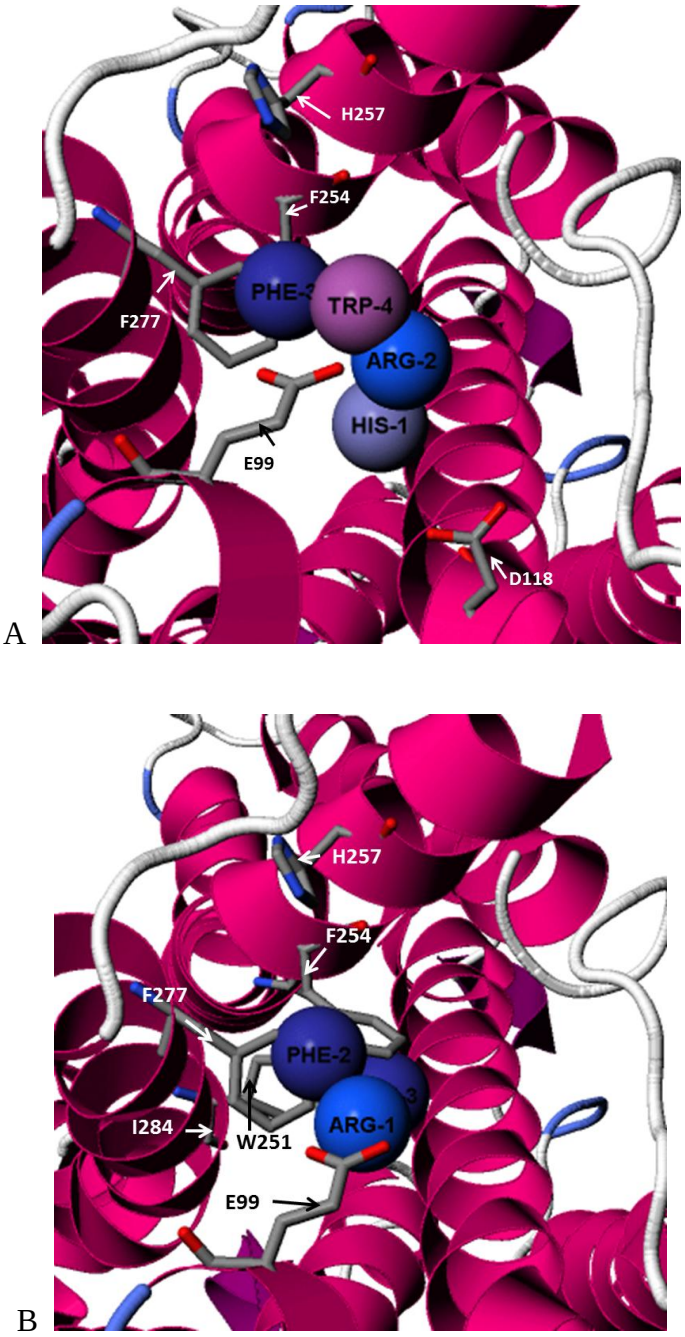


Figure 5.18. Computer prediction from Pepsite showing the position of the A) HFRW fragment and B) RFF fragment from ASIP in the binding pocket of the MC1R-wt

The data gathered from the three sources above (MCR functional data, crystallographic data and computer predictions) are presented in appendix 5, which shows the residues predicted to

be in the binding pocket, and likely to be involved in ligand-receptor interaction in the MC1R-wt and MC1RΔ24. Each single piece of evidence suggesting that a particular residue is in the MC1R binding pocket is represented as a cross in the table. However, not all data points were considered to be equal; some data provide stronger evidence than others. For example, a particular residue of the human MC1R may be both conserved with that of the MC1R-wt and also found to be significantly involved in ligand binding. Here, this is considered to be stronger evidence that the residue in question may be involved in the MC1R-wt binding pocket than a computer prediction based on a homology model. In order to account for the different levels of relevance of the data, a scoring system was devised (see table 5.1)

Table 5.1 Scoring system for predicting residues in the binding site of the MC1R

The score shows the number of points awarded for a particular piece of evidence ranging from 1 point for a residue predicted by computer to 7 for a residue which is conserved and also shown to be involved in binding from MC1R functional data.

Score	Justification
7	MC1R functional data (conserved residue)
6	MC1R functional data (not conserved)
5	MC4R functional data (conserved residue)
4	MC4R functional data (not conserved)
3	A2AR crystallography data (conserved residue)
2	A2AR crystallography data (not conserved residue)
1	Computer prediction

If a particular residue had a score of 15 or above, from appendix 5, it was considered likely that this residue would be involved in the ligand binding pocket of the MC1R. Figure 5.19 shows all residues with a score of 15 or above represented on a diagram of the MC1R-wt and MC1RΔ24. These residues are subsequently considered to be "consensus" residues.

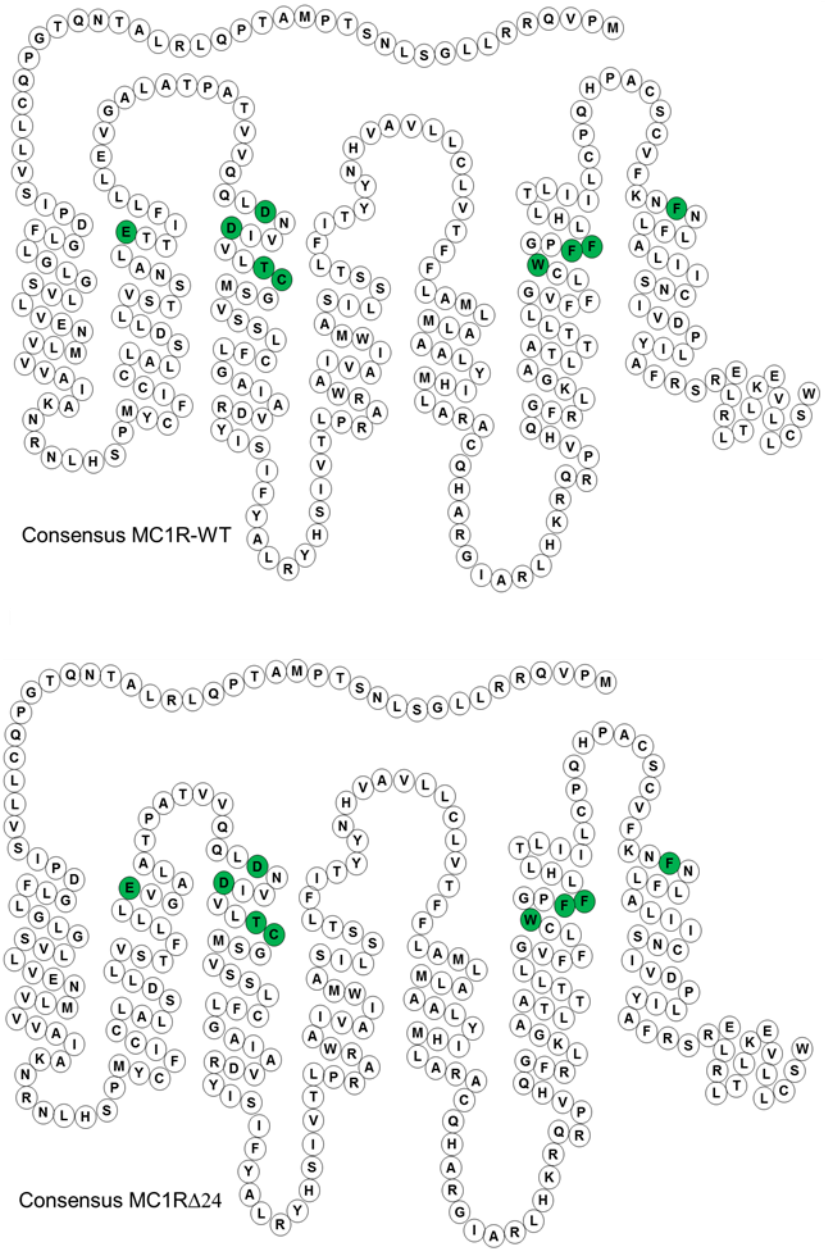


Figure 5.19. Schematic representations of the MC1R-wt and MC1RΔ24 showing amino acids predicted to be in the binding pocket

Circles coloured green are predicted to be involved in the binding pocket.

5.2.5 Regions of the binding pockets

Figure 5.20 shows the consensus of residues highlighted on a computer model of the MC1R.

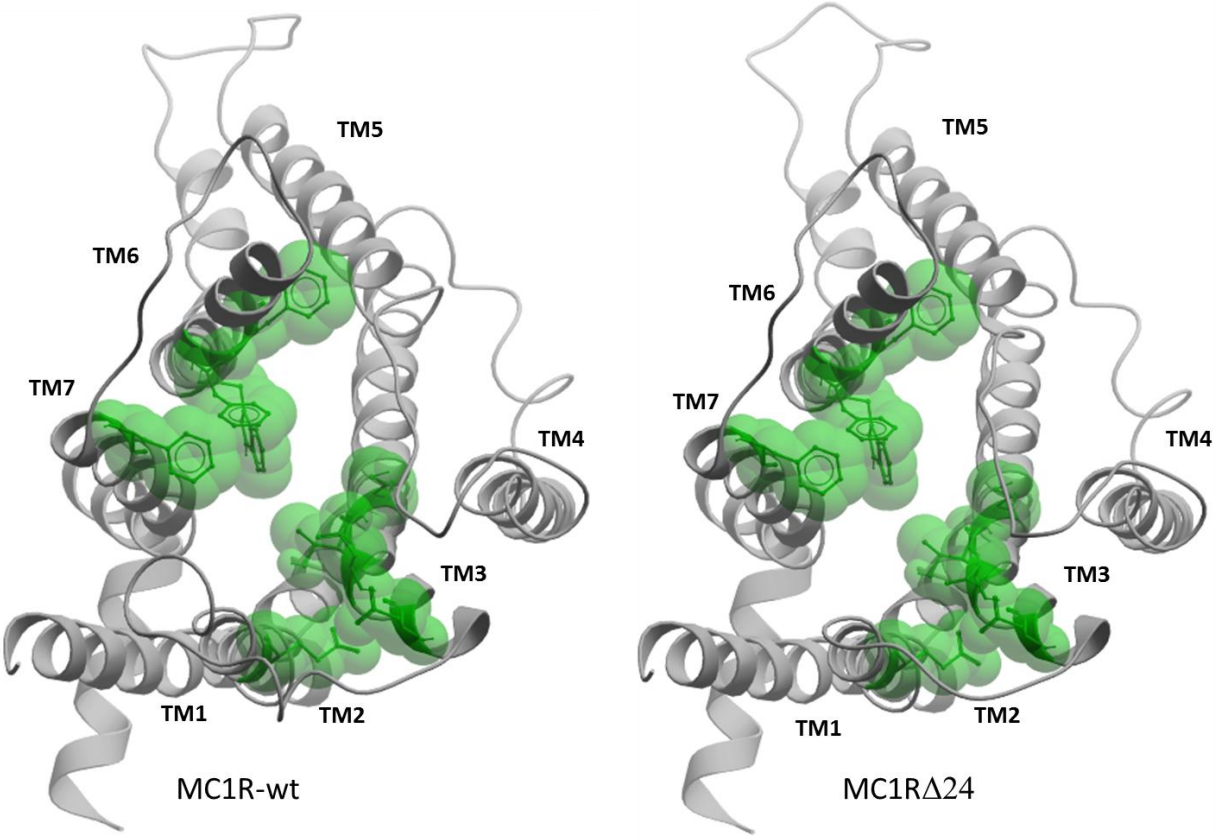


Figure 5.20. Extracellular view of the MC1R-wt and MC1RΔ24 showing the positions of amino acids predicted to be in the binding pocket by experimental data and computer programmes

Ribbon diagram representation of the MC1R-wt and MC1RΔ24 with amino acids predicted to be in the binding pocket shown as space-filling models coloured in cyan. TM= transmembrane helix. The MC1R is visualised with Molsoft.

Figure 5.21 shows the two distinct regions of the receptors' binding pockets, showing an acidic region (red) and an aromatic region (blue).

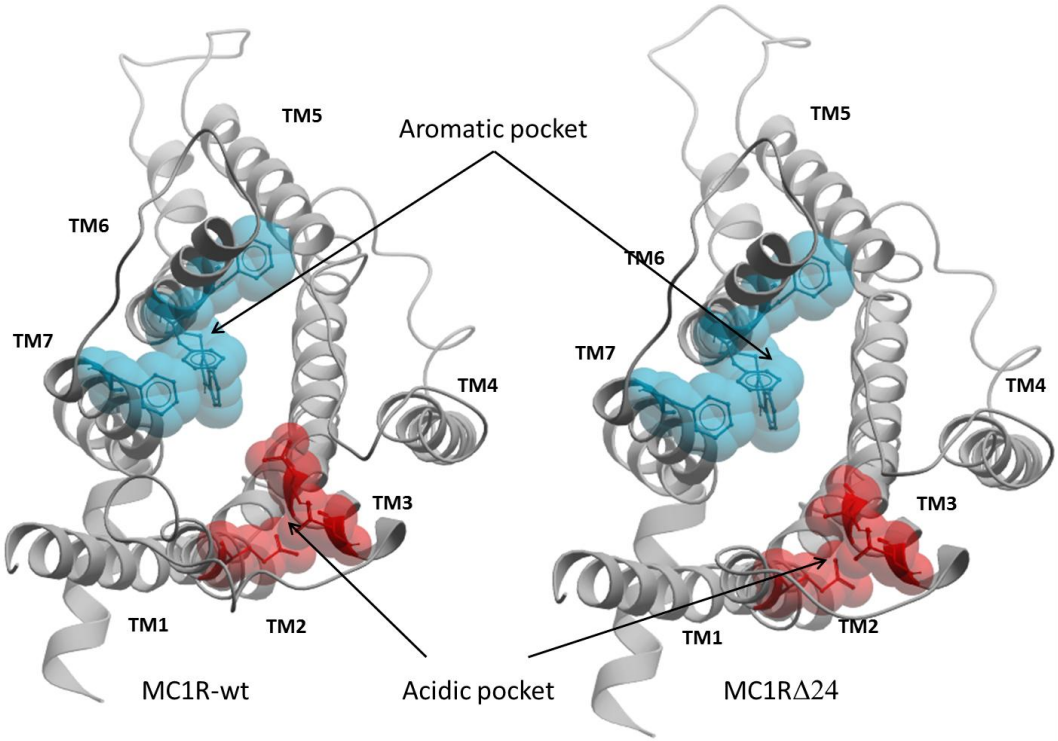


Figure 5.21. Extracellular view of MC1R-wt and MC1RΔ24 showing different regions of the binding pocket predicted by experimental data and computer programmes

Ribbon diagram representations of the MC1R-wt and MC1RΔ24 with amino acids of the binding pocket, predicted by experimental data and computer programmes, as space-filling models coloured red to show the acidic region and blue to show aromatic residues. The MC1R molecules are visualised with Molsoft.

5.2.6 Structure of the alpha-melanocyte stimulating hormone

The structure for the α-MSH was determined using the Pepfold web server. The sequence however, is short and likely to be highly flexible (Orry and Abagyan, 2012). Figure 5.22 shows three possible predicted structures of the α-MSH showing side chains.

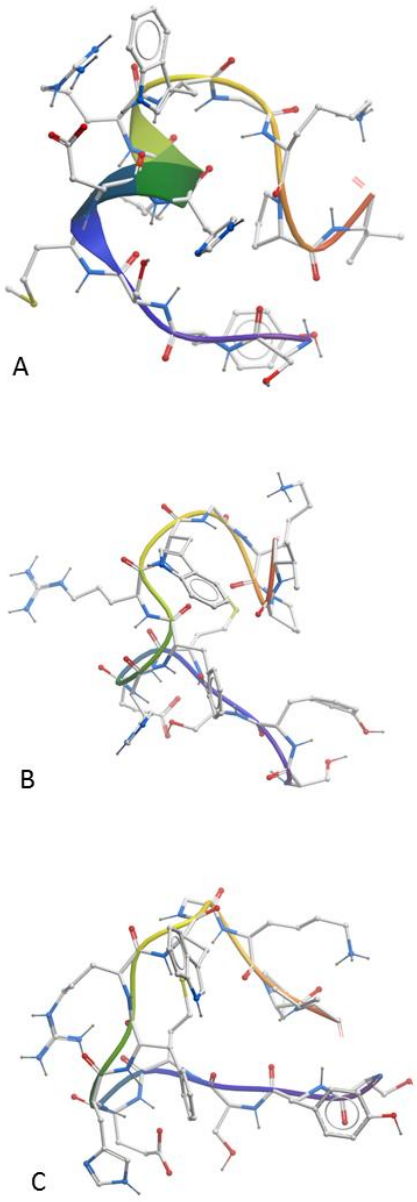


Figure 5.22. Computer models of the predicted structure of α-MSH using the web server Pepfold

Three different models of the α-MSH peptide predicted by Pepfold. Atoms are shown as sticks with a smoothed backbone which is coloured using the rainbow spectrum from blue at the N- terminus to red at the C-terminus. The thickened backbone in A is an alpha helix.

5.2.7 Structure of the agouti signalling protein

The partial structure of human ASIP has been determined by NMR (residues 80-132) and here the squirrel ASIP sequence is matched to this by Phyre2 and I-TASSER (PBD ID: 1Y7J). Figure 5.23 shows ASIP with the highly conserved RFF side chains highlighted and three disulphide bridges shown as white and yellow sticks. ASIP structure has an unusual inhibitor cysteine knot fold motif which forms a topological circle created by three disulphide bonds which hold the molecule in a highly constrained position (McNulty et al., 2005).

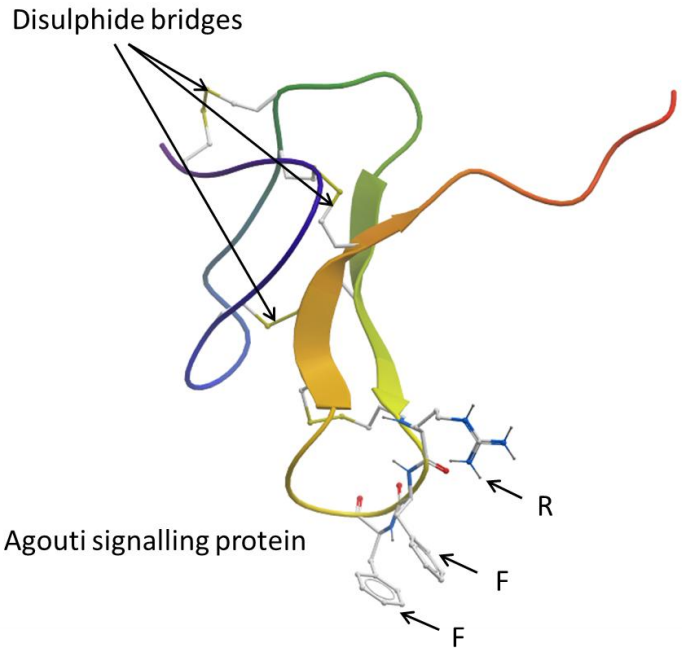


Figure 5.23. Computer model prediction of agouti signalling protein

Agouti signalling protein (residues 80-132), showing the RFF sequence motif highlighted with side chains as stick representations. Disulphide bridges are shown as white and yellow sticks. The backbone is coloured using the rainbow spectrum from blue at the N- terminus to red at the C-terminus. The model is visualised with Molsoft.

5.2.8 Molecular Docking

In order to gain an understanding into how ligands may bind to the receptors, computer docking was carried out. Docking programmes aim to predict how a ligand will interact with a receptor within the constraints of a receptor binding site and aim to estimate the strength of that binding. Here the programme *CLC Discovery* was used to investigate how the MC1R-wt and MC1RΔ24 interact with α-MSH and ASIP. The three different α-MSH models, A, B and C shown in figure 5.22 were used and a model of the α-MSH fragment HFRW was also used, making a total of eight combinations. Each combination was then repeated eight times. Docking was also carried out with ASIP but no energetically favourable binding was found. Results are presented in table 5.2. The scoring function used by *CLC Discovery* is complex and does not simply give an energy value. The score takes into account a number of factors, where hydrogen bond interactions, ionic interactions and non-polar interactions are rewarded and non-polar to polar contacts and repulsive contacts are penalised. Therefore the score does not have units but gives an overall indication of the predicted docking interaction. Results shown in table 5.2 show that for both receptors, there is a good binding affinity predicted for this sequence and figure 5.24 show the predicted conformations. However, as the figure 5.24 illustrates, the conformation of α-MSH interacting with the wildtype and mutant were completely different.

Table 5.2. Docking scores from *CLC Discovery* for MC1R-wt and MC1RΔ24 interacting with three different models of α-MSH and the HFRW fragment of α-MSH

Scores are ordered with the most favourable score (most negative) at the top. The scores have no units and are derived from calculations where hydrogen bond interactions, ionic interactions and non-polar interactions are rewarded and non-polar to polar contacts and repulsive contacts are penalised.

	α-MSH Model A		α-MSH Model B		α-MSH Model C		HFRW	
	MC1R-wt	MC1RΔ24	MC1R-wt	MC1RΔ24	MC1R-wt	MC1RΔ24	MC1R- wt	MC1RΔ24
1	-43	-21	-22	7	-16	-19	-91	-96
2	-9	4	22	18	21	21	-82	-87
3	4	24	38	39	41	32	-79	-86
4	14	29	43	49	49	43	-78	-85
5	15	34	51	56	68	60	-75	-84
6	16	38	58	64	77	73	-73	-81
7	44	42	78	78	85	83	-73	-79
8	57	47	99	88	108	108	-70	-78
Ave.	12	25	46	50	54	50	-78	-85

The lowest scores for each model were negative with the exception of MC1RΔ24 receptor with model B. This negative number indicates that binding is energetically favourable and figures 5.24 and 5.25 show the conformations predicted. However, the docking scores averaged over the eight simulations, for the three α-MSH models with both MC1R-wt and MC1RΔ24 receptors were all positive. A positive score is considered by *CLC Discovery* to predict that there would be no favourable binding.

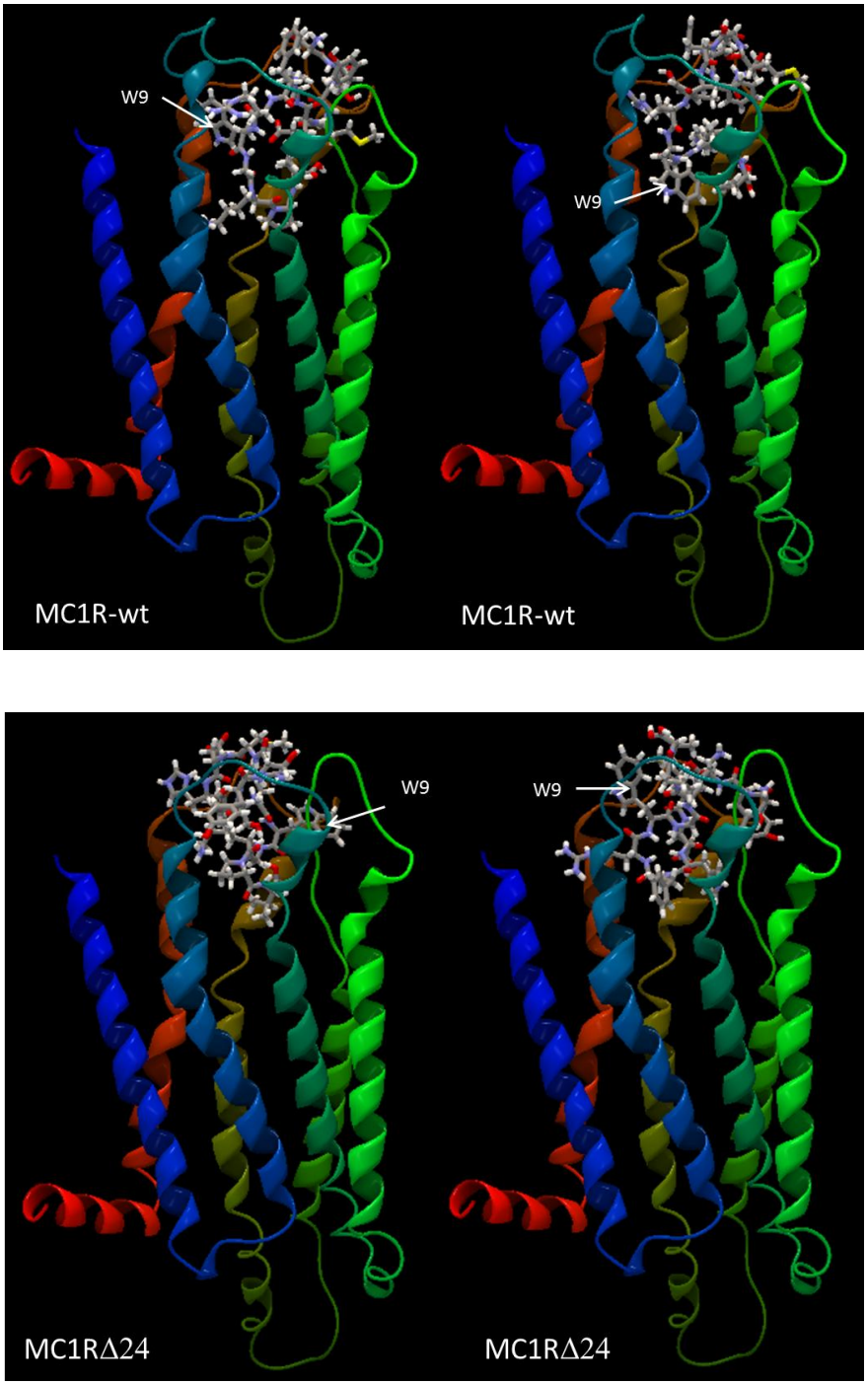


Figure 5.24. Orientations of α-MSH in the binding site of the MC1R-wt and MC1RΔ24 receptors as predicted by *CLC Discovery*

The MC1R in each case is a ribbon representation and α-MSH is a stick representation. The tryptophan residue (W9) of α-MSH is highlighted in all cases.

Figure 5.25 shows how the α-MSH fragment HFRW was docked into each receptor. Only one representative of each variant is shown in this case.

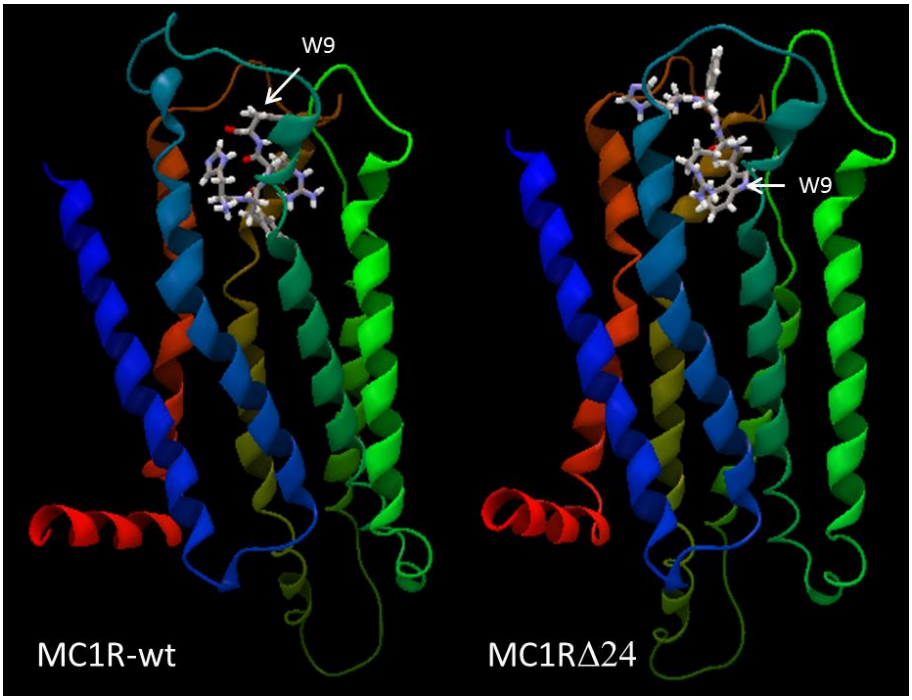


Figure 5.25. Orientations of the HFRW fragment of α-MSH in the binding site of the MC1R-wt and MC1RΔ24 receptors as predicted by *CLC Discovery*

The MC1R in each case is a ribbon representation and HFRW of α-MSH is a stick representation. The tryptophan residue (W9) of α-MSH is highlighted in both cases.

5.3 Discussion

5.3.1 Homology modelling of the melanocortin-1 receptor

Predicting the native, three dimensional structure of a protein from its sequence is a difficult problem and has led to the development of many powerful computer techniques (Xiang, 2006; Orry and Abagyan, 2012). There are no crystal structures available for any melanocortin receptors and the structures presented here are based on homology modelling. At present, homology modelling provides the best method available for predicting protein structures where no experimental data is available, and it is a great deal more powerful than *ab initio* predictions. Homology modelling relies on the observation that there are limited ways that proteins fold in nature and sequences are matched as closely as possible to proteins of known structure (Kaczanowski and Zielenkiewicz, 2010). Amino acid side chains are then "threaded" onto the template backbone to create the predicted three dimensional model. This model is then iteratively improved by excision and reassembly (Xiang, 2006). As more and more protein structures are solved, homology modelling will become more and more accurate. Here, two independent computer systems were used to arrive at the predicted models and both systems matched the MC1R to the A2AR with high confidence. High confidence in this case means that the A2AR is the closest match, and indeed, this is consistent with the evolutionary tree of the GPCR superfamily (figure 1.28) where receptors with 35% homology with the A2AR are highlighted in yellow. Homology models have limitations, and differences between the template and the MC1R will inevitably affect the accuracy of the predicted structure. Here, for example, three residues that are thought to be of particular importance in the MC1R are E91, D114 and D118 (MC1R-wt numbering), and it is thought that these residues may contribute to a repulsion between TM2 and TM3 (Lu, Vage and Cone, 1998). The alignment in figure 5.3 and 5.4 shows that the templates on which the

MC1R models are built do not have these residues. As these residues are not the same in the template, it is possible that the predicted relative positions of TM2 and TM3 are incorrect. The models presented here are the best available approximations to date of the three dimensional structure of the MC1R, and while the limitations of these models are fully acknowledged, the following discussions assume that they are broadly correct.

5.3.2 Structural effects of the 24 bp deletion and interactions between agouti signalling protein and extracellular loop 1

Phyre2 and I-TASSER models both independently predict that the deletion of the SNALETTI sequence from the wildtype receptor causes the FLLLEVGALA (95-104 MC1R-wt numbering) sequence of ECL1 to be incorporated into TM2 in the mutant. This change causes the E99 from the wildtype to be relocated to position 91 in the mutant, so that both receptors maintain an acidic region at this position in the binding pocket, which is important for ligand binding (figure 5.21) (Lu, Vage and Cone, 1998). This change also results in a shorter ECL1 in the mutant. ECL1 of the MC1R-wt (13 residues) has a β strand, but this strand is not present in the mutant ECL1 (4 residues). Given that ECL1 is essential for ASIP binding, and that the C-terminus loop of ASIP is involved in hydrophobic interactions with this ECL (Patel et al., 2010), it seems likely that ASIP binding to the mutant may be affected. Figure 5.26 shows how the C-terminus loop of ASIP is closer to ECL1 of the MC1R-wt than that of MC1RΔ24. It is predicted that ASIP will not bind to the MC1RΔ24 and this possibility is explored further in functional experiments presented in chapter 6.

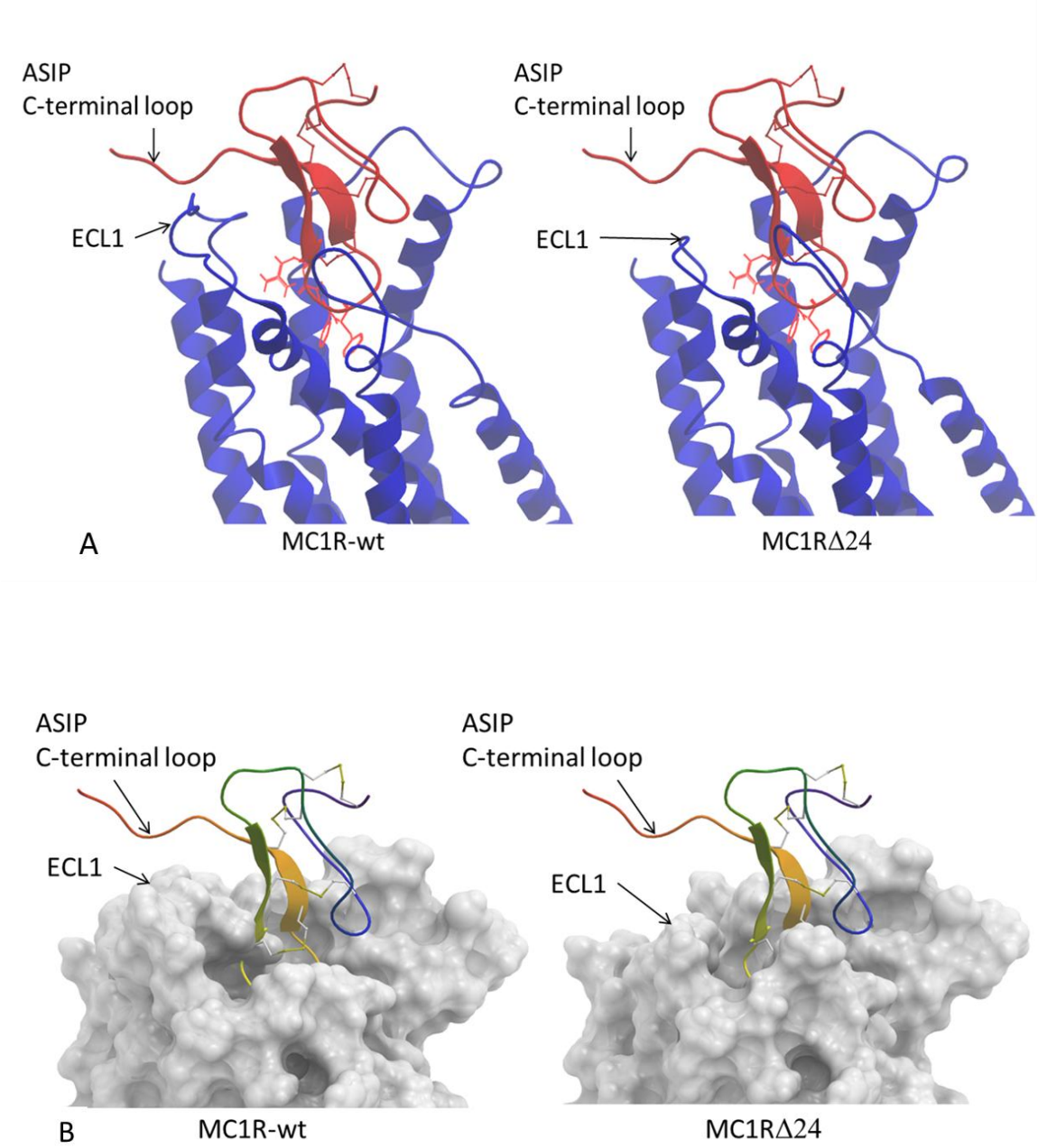


Figure 5.26. Computer models of MC1R-wt and MC1RΔ24 with agouti signalling protein

A) Ribbon diagram representation of MC1R-wt and MC1RΔ24 with ASIP showing the RFF motif in the active loop of agouti signalling protein (ASIP) as a stick representation and B) showing the ASIP C-terminal loop and the MC1R molecular surface. ASIP is shown as a smoothed backbone with rainbow colouring from blue at the N-terminus to red at the C-terminus.

5.3.3 Structural changes to transmembrane helix two and transmembrane helix three in the MC1RΔ24

The eight amino acid deletion appears to cause substantial overall structural changes to the receptor. The models show that the wildtype receptor has a shorter TM2 with 25 residues than the mutant receptor which has 27 residues (figure 5.8). These large changes are likely to lead to different intramolecular interactions in each case. The data from Socket, however, show identical "knobs into holes" interactions for both the wildtype and MC1RΔ24 receptors. Socket data also show that the highest number of interactions between any two helices is between TM2 and TM3 (figure 5.11). It has already been noted that the deletion abolishes the hydrogen bond predicted to be between N88 of TM2 and S124 of TM3 and it seems likely that there will be other differences in packing of these helices between the wildtype and MC1RΔ24 receptors. Considering that the MC1RΔ24 has two more residues in TM2, it seems likely that there would be more interactions between TM2 and TM3 in the MC1RΔ24 compared to the wildtype receptor. Given the association between TM2 and TM3, it seems likely that the large architectural changes to TM2 caused by the deletion would have an impact on TM3. TM3 acts a structural hub and structural analysis of crystallized GPCRs show that, during activation it undergoes a 2-14 Å shift towards the extracellular side of the receptor (Venkatakrishnan et al., 2013). It may be hypothesised that the TM2 differences and ECL1 shortening could lead to an upward shift of the TM2-TM3 bundle towards the extracellular side thus mimicking the upward movement observed in active receptors. Based on this observation, it is predicted that the MC1RΔ24 is a constitutively active receptor and this is prediction is explored in chapter 6. This change in position of TM2 and TM3 is not, however, apparent in the homology models and the G-protein binding pockets appear to be largely the same for both variants of the receptor. This is likely to be a reflection on the

limitations of the modelling in which both the wildtype and mutant MC1R share a single template which will tend to minimise differences between them. It is anticipated that, in reality, there would be differences at the intracellular side of the molecule caused by the deletion which could affect G-protein activity.

5.3.4 Residues of the melanocortin-1 receptor binding pockets

The results of investigations into the binding pockets of the MC1R show that a range of different residues may be involved with ligand binding. There was a general discrepancy between the residues predicted to be in the binding pockets by Fpocket, Pepsite and those established from experimental evidence. Specifically, the acidic pocket created by the E91, D114 and D118 residues was not consistently recognised by the computer programmes. This may be due to the misplacing of side chains in the model due to limited availability of suitable template structures for homology modelling. However, the data from Pepsite gave a useful insight into the specific interactions predicted between the pharmacophore HFRW of α-MSH and RFF of ASIP and the MC1R, which were in general agreement with data collected from mutagenesis experiments. The results in appendix 5 show that there is strong evidence to support the prediction that the following residues are involved in ligand binding: E91, D114, D118, T121, C122, W251, F254 and F277 (MC1R-wt numbering) and these residues are conserved in the MC1RΔ24 variant (figure 5.19). It has been suggested by Yang (2002) that melanocortin peptides use conserved residues for ligand binding (Yang et al., 2002). This is consistent with the prediction that conserved residues from human MC1R and MC4R, found to be involved in binding, would also be the case for the squirrel MC1R. Taken together, the data collected from computer predictions and experimental evidence from

MCRs and A2AR allow a consensus to be reached on which residues are likely to be involved in ligand binding and these residues are present in both MC1R variants (figure 5.19).

5.3.5 Transmembrane helices involved in the binding pocket

TM1 does not contribute to the binding pocket, as shown in figures 5.18 and 5.19. The consensus in figure 5.19 predicts that TM2, 3, 6 and 7 all contribute to the binding pocket. All experimental data indicates that TM6 contributes more residues than any other TM and that these are generally aromatic residues. There is general agreement that TM6 undergoes a tilting movement during activation of GPCRs where the intracellular end moves outward and the extracellular end moves inwards towards TM3 (Farrens et al., 1996; Gether et al., 1997). This movement is triggered by ligand binding. Like other MCRs, the MC1R of the squirrel has a short ECL2 compared to other GPCRs. This short ECL is thought to allow free inward movement of TM6, which is thought to account for MCRs having a high degree of constitutive activity (Holst and Schwartz, 2003). Furthermore, the disulphide bridge formed in ECL3 is thought to hold TM6 in towards TM3 and agonist binding would stabilise this active conformation. It is hypothesised that inverse agonists function to hold TM6 away from TM3 (Holst and Schwartz, 2003). TM6 contributes to a hydrophobic and aromatic region of the binding pocket where the His, Phe and Trp residues of the HFRW pharmacophore and the double Phe residues of the RFF motif are thought to interact (see figure 5.18). This region of the receptor appears to be the same for MC1R-wt and MC1RΔ24; however, the interactions between TM6 and TM3 are likely to be subtly altered due to the substantial changes to TM2 in the mutant. TM2 and TM3 both have residues which have been shown to be important for

ligand binding in numerous studies (Holst and Schwartz, 2003). The E91, D114 and D118 (MC1R-wt numbering) contribute to an acidic pocket, as shown in figures 5.15 and 5.21. Although both MC1R-wt and MC1RΔ24 have this acidic region, the specific side chains and their orientations which may interact with ligands are likely to be different in each case.

5.3.6 Molecular docking

The docking results from *CLC Discovery* indicate that α-MSH binds with greater affinity to the wildtype than to the mutant receptor, but the average scores for both receptors suggest no favourable binding. The best α-MSH model appears to be model A (figure 5.22) which appears to bind with greater affinity than the other two models. The HFRW fragment binds well to both receptors but with a slightly higher affinity for the mutant receptor. It should be noted that binding affinity is not an indication that the ligand will activate the receptor. There are many cases where binding affinity has been shown experimentally to be strong but where the ligand acts as an antagonist, holding the receptor in the inactive conformation. Examples include the artificial ligand SHU9119 with MC4R (Lee et al., 2010) and the loss of function MC1R mutants I40T, V122M, R151C, R162P and L93R (Más et al., 2002; Sánchez-Más et al., 2004). *CLC Discovery* did not find any favourable docking solutions for ASIP. This is not surprising as it is known that ASIP binding requires the accessory protein attractin in order to function with the MC1R (He et al., 2001).

The docking programme *CLC Discovery* was selected as it was able to take into account the flexibility of the ligand. Computer models of α-MSH predict that the peptide has 49 rotatable

bonds and *CLC Discovery* returned many results with the ligands in various orientations and conformations as demonstrated by the Trp residue in different positions in figures 5.24 and 5.25. Indeed, it might be the case that many different orientations and conformations of the ligand allow energetically favourable binding but that only one or a few orientations lead to activation. Although both the wildtype and mutant receptors have the same conserved residues in the binding pocket, there are likely to be many subtle differences between the two variants which may affect ligand binding. Given the flexible nature of both the ligands and receptors it seems likely that an induced fit would lead to different conformations which could affect signalling. The full extent of the flexibility of α-MSH is not known and it is possible that this peptide is a disordered structure, lacking a distinct fold in the unbound state (Orry and Abagyan, 2012). It is also possible that on meeting, the peptide and receptor both change in accordance with the induced fit model or alternatively go through conformational sampling (Orry and Abagyan, 2012). The differing binding pockets found in the A2AR shown in figure 5.16 are consistent with the induced fit model. Experimental evidence from A2AR with NECA and caffeine show that the same receptor can produce binding pockets of different shapes and sizes according to the ligand (Lebon et al., 2011). Indeed, results from binding studies of MC4R and ASIP have indicated that receptors have distinct states for different ligands. This is contrary to a simple two-state conformational selection model where the ligand either binds to the active or the inactive receptor conformation (Chai et al., 2003; Chai et al., 2005). Thus, experimental evidence from MC4R supports the induced fit model of ligand-receptor interaction, where both receptor and ligand are flexible and able to change conformation on binding.

Generally, the data obtained from docking simulations here should not be taken as an accurate prediction of the interactions that would take place *in vivo.* For docking to give meaningful results, an accurate starting model is required and both the ligand and receptor would need to be flexible. Although homology modelling gives the best prediction possible, it does not provide an accurate enough model for such fine detailed interactions required for docking (Xiang, 2006). Furthermore, extracellular loops are likely to be involved with ligand binding and the structures of these ECLs are particularly hard to predict, even from crystal structures. In the absence of experimental structural data on positioning of the ligands in the MC1R, it is therefore not possible to evaluate the returned data in the usual manner using RMSD values to compare crystal structure data with docking data. Unfortunately, *CLC Discovery* is not able to simulate a flexible receptor, which could account for the unfavourable binding of α-MSH, and therefore it is not a realistic model. Although there are too many degrees of freedom in this system and limitations in the model and computer programmes, overall the data does show a good geometrical and chemical complementarity between α-MSH and the MC1R and some plausible interactions are suggested. Figure 5.27 shows how α-MSH fills the proposed binding site with an overall good fit.

.

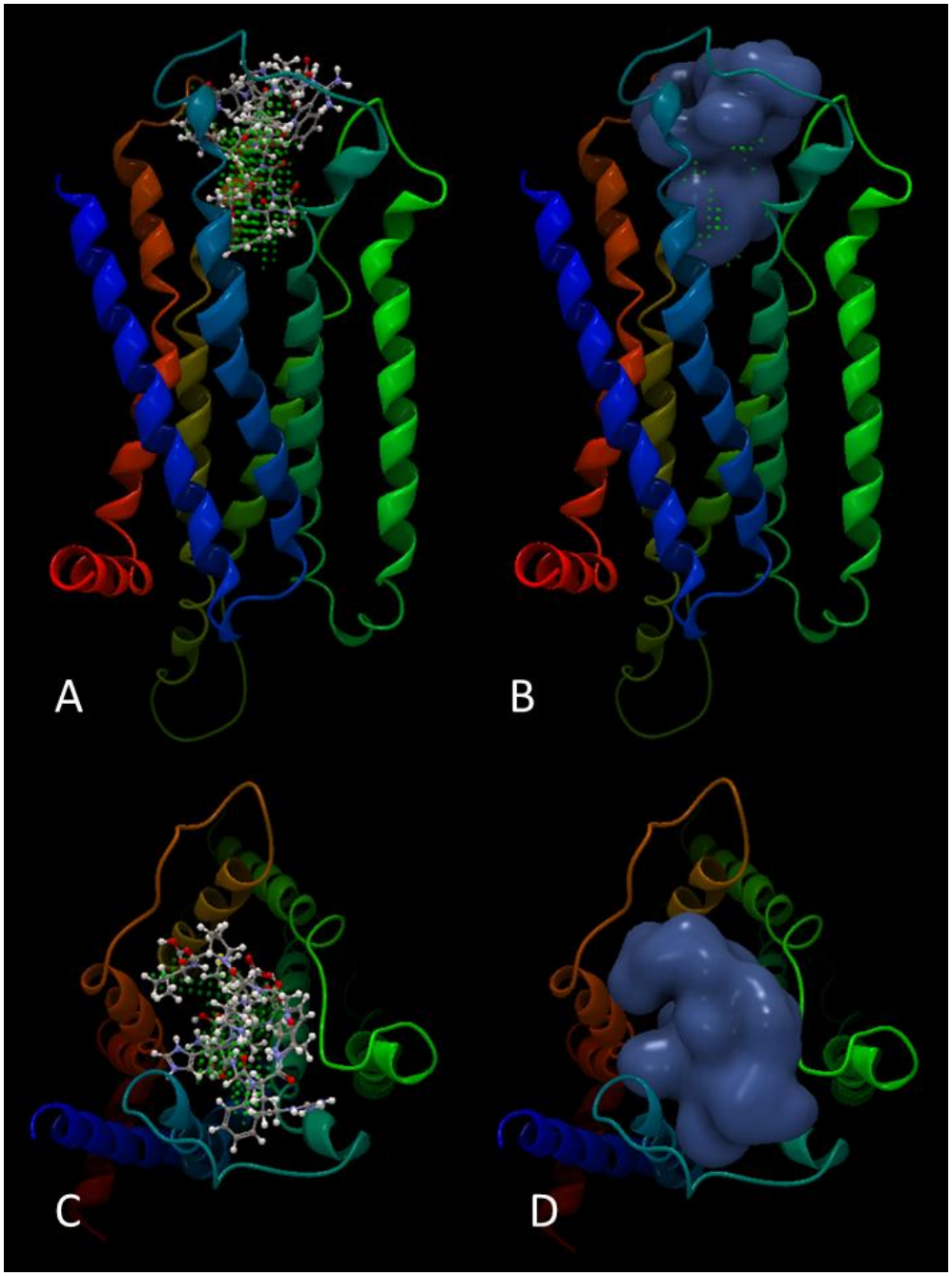


Figure 5.27. Computer models of MC1R-wt docked with α-MSH

Ribbon diagram representations of the MC1R-wt showing A) side view of the binding pocket shown as green spots and stick representation of α-MSH, B) side view showing α-MSH molecular surfaces, C) and D) extracellular view with α-MSH as stick and molecular surface representations respectively.

5.3.7 HFRW motif interactions with the melanocortin-1 receptor

Extensive mutagenesis experiments have led to a good understanding of the interactions between the MC1R and its ligands. Each residue of the HFRW pharmacophore from α-MSH has been analysed by alanine scanning studies which show that the His is important for both potency and binding affinity (Holder and Haskell-Luevano, 2004). The data suggest that this residue plays a role in stabilising the ligand-receptor complex and that the imidazole ring allows conformational change to take place within the binding site. The Phe is also thought to contribute to a stable complex and is likely to interact with the aromatic residues in the binding pocket (Holder and Haskell-Luevano, 2004). The Arg is considered to be essential for receptor activation as it forms an ionic interaction with the acidic pocket in the binding site (Holder and Haskell-Luevano, 2004). It has also been suggested that the two terminal NHs of the Arg side chain are essential for interactions rather than the positive charge. The Trp is essential for aromatic interactions with the receptor (Holder and Haskell-Luevano, 2004). Similar findings have been observed for the RFF motif from ASIP where the Arg is thought to interact with the acidic pocket and the Phe residues both form aromatic interactions in the pocket. It has been observed that the spatial orientation of the aromatic groups is more important than the charge-to-charge interactions (Lee et al., 2010). Indeed, it has been shown that compounds with no basic residue are capable of generating agonist activity in the receptor (Holder and Haskell-Luevano, 2004). A number of different compounds have been investigated for the activity and binding to the MC1R with many acting as potent agonists and others as antagonists. Taken together, it may be predicted that the α-MSH will bind to and activate both the MC1R-wt and MC1RΔ24 in similar but not identical ways with the Arg from the HFRW and RFF interacting with the acidic pocket and the aromatic residues interacting with the aromatic residues of TM6.

5.4 Conclusion

Computer models predicted that the eight amino acid deletion in the MC1RΔ24 leads to a receptor with a shorter ECL1 than the wildtype. ECL1 has been found to be critical for ASIP binding and the shortening in the mutant is predicted to affect the receptor's response to ASIP. The models predict that TM2 is lengthened in the mutant and that this may alter the packing between TM2 and TM3. The altered interactions between TM2 and TM3 are predicted to lead to a constitutively active receptor. Both MC1R-wt and MC1RΔ24 have acidic and aromatic regions in the extracellular binding pockets but the deletion is likely to have made subtle changes which could affect ligand binding and function. The functional effects of the eight amino acid deletion in the MC1R are explored in the next chapter where the MC1R-wt and MC1RΔ24 responses to ligands are assessed *in vitro*.

Chapter Six

Functional Characterisation of the MC1R-wt and MC1RΔ24

6.1 Introduction

A number of species have mutations in the *MC1R* gene which are associated with melanism, as already discussed (Majerus and Mundy, 2003). MC1R activity is coupled to intracellular cAMP production via G_S-protein and adenylate cyclase, as described in the introduction. Functional studies have revealed that some of these mutations cause the receptor to be constitutively active, for example the E92K mutation in mice (Robbins et al., 1993), chickens (Ling et al., 2003), and lemurs (Haitina et al., 2007) and the C125R mutation in foxes (Våge et al., 1997a) and the M73K mutation in sheep (Våge et al., 1999). To date, no deletions in the *MC1R*, associated with melanism, have been functionally characterised. The aim of this part of the study was to functionally characterise the MC1R-wt and MC1RΔ24 by expressing the receptors in HEK293T cells and stimulating the transfected cells with α-MSH and ASIP, then measuring intracellular levels of cAMP. The initial working hypothesis was that the eight amino acid deletion in the MC1RΔ24 would lead to a constitutively active receptor and the second hypothesis was that the MC1RΔ24 would be unresponsive to the antagonist ASIP.

The methods used in this section included PCR, gel electrophoresis, DNA sequencing, cloning, cell culture, transfection of cells, SDS-PAGE, Western blotting, flow cytometry, protein assay, fluorescence microscopy, site-directed mutagenesis and measurement of cAMP by ELISA. Data was analysed using Student's t-test.

6.2 Results

6.2.1 DNA constructs

In order to characterise the MC1R-wt and MC1RΔ24, the genes were amplified by PCR from genomic DNA (figure 6.1). Initially, the genes were amplified with the MC1Rexpf1 forward primer, which included an initial Kozak sequence and start codon, and the MC1Rexpstopr1 reverse primer which included a stop codon. This PCR product was cloned into a pcDNA3.1/V5-His TOPO® TA Expression vector (Invitrogen) so that the whole *MC1R* gene would be translated with no extra tags. However, it was not possible to detect the MC1R proteins in transfected cells using any MC1R antibodies commercially available. Therefore, instead, the genes were amplified using the MC1Rexpf1, which had the initial Kozak sequence, and the MC1Rexpr1 reverse primer, which removed the stop codon and added a G at the end of the gene, in order to be in frame for expression of a C-terminal GFP tag. The PCR products from this reaction were cloned into the CT-GFP Fusion TOPO® TA Expression vector (Invitrogen). The MC1R-GFP fusion protein was successfully detected from transfected cell using this method. It has been demonstrated that C-terminal tagging of the MC1R does not affect function (Sánchez-Laorden et al., 2006).

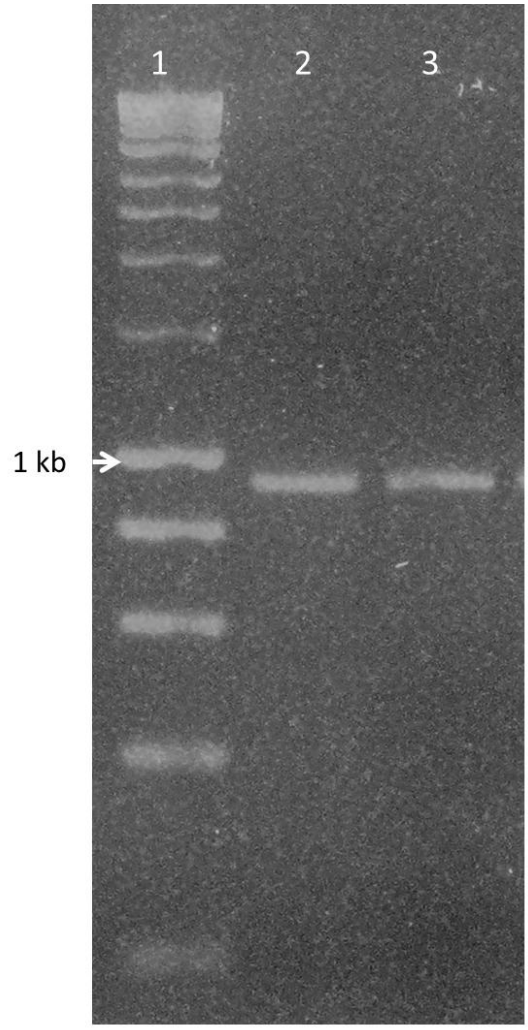


Figure 6.1. PCR amplification of the *melanocortin-1 receptor* gene from the grey squirrel *Sciurus carolinensis*

A representative agarose gel electrophoresis (1.3%) of PCR amplification products visualised with ethidium bromide and illuminated with UV light. Lane 1, Hyperladder 1 (Bioline) DNA ladder; lanes 2 and 3 show a 950 bp (approximately) product of the MC1R-wt and MC1RΔ24 genes, respectively, amplified with the primers MC1Rexpf1 and MC1Rexpr1.

The DNA constructs were cloned as described in Materials and Methods (chapter 2).

Plasmids were extracted from bacteria and run on an agarose gel to be visualised (figure 6.2).

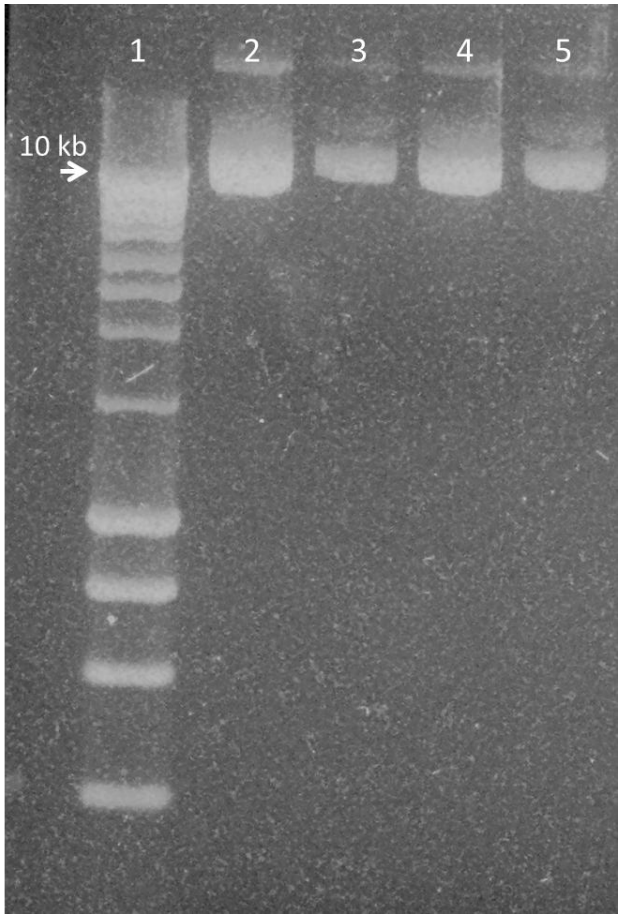


Figure 6.2. MC1R-wt and MC1RΔ24 GFP fusion TOPO plasmids

A representative agarose gel electrophoresis (1.3%) of plasmids visualised with ethidium bromide and illuminated with UV light. Lane 1, Hyperladder 1 (Bioline) DNA ladder; lanes 2 and 3 show GFP fusion TOPO plasmids with MC1R-wt (1µl and 0.5µl respectively) and lanes 4 and 5 show GFP fusion TOPO plasmids with MC1RΔ24 (1µl and 0.5µl respectively). Bands of each plasmid are approximately 10 kb.

6.2.2 Cell transfection

In this study, it was not possible to use melanocytes obtained from the grey squirrel to study the MC1R. When melanocytes cannot be used directly to study MC1R function, HEK293T cells are often used to express the gene in a heterologous way. These cells give robust results on MC1R function as there is no endogenous expression of the MC1R (Mountjoy et al.,

1992). Therefore HEK293T cells were used to transiently express the MC1R variants in this study. Expression of the MC1R variants was confirmed by direct visualisation of the GFP with a fluorescence microscope. GFP could be seen in the cells, as shown in figure 6.3. Transfected cells showed a diffuse fluorescence throughout the cell which is thought to be GFP in the membranes. Cells also showed a region of concentrated fluorescence which is thought to be GFP in the ER (see figure 6.3 C).

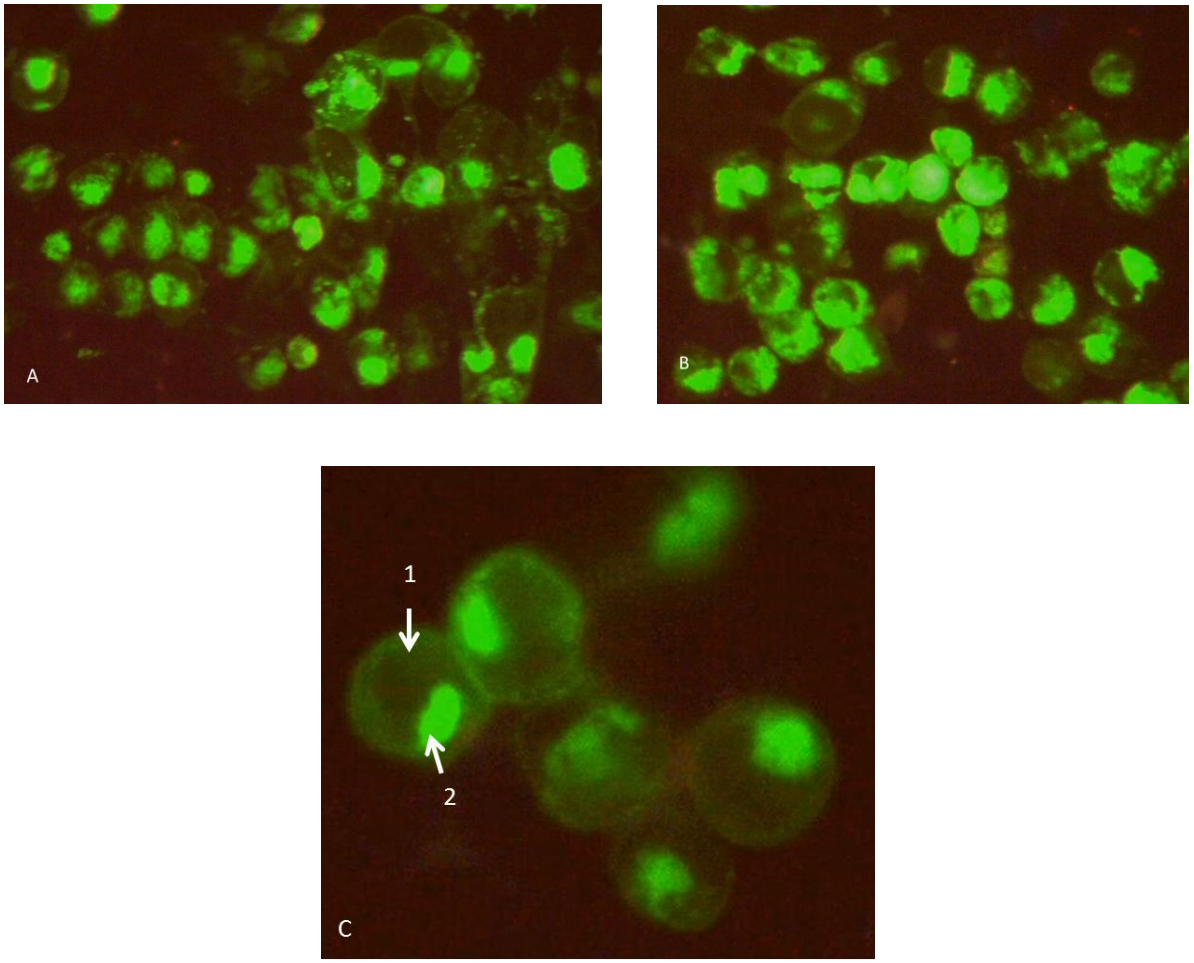


Figure 6.3. HEK293T cells expressing MC1R-GFP fusion proteins

Transiently transfected cells viewed under a fluorescence microscope showing A) cells transfected with the MC1R-wt fused with GFP and B) cells transfected with the MC1RΔ24 fused with GFP and C) close up view of cells expressing MC1R-wt-GFP fusion protein. C) 1 shows diffuse fluorescence which is thought to be MC1R-wt GFP fusion protein in the cellular membrane and C) 2 shows MC1R-wt GFP fusion protein thought to be in the endoplasmic reticulum.

6.2.3 Western blot

Unfortunately, it was not possible to directly detect the MC1R of the squirrel by Western blotting. After numerous attempts, it was concluded that available antibodies could not bind to the squirrel protein. When the *MC1R* variants were cloned into the CT-GFP Fusion

TOPO® TA Expression vector (Invitrogen), an antibody to GFP (GFP rabbit serum polyclonal antibody Invitrogen) was used instead. The MC1R-GFP fusion protein created a product of approximately 64 kDa, where MC1R is typically 23-35 kDa and GFP adds at least 27 kDa to the protein (Beaumont et al., 2005). Blots were probed with the anti-GFP to detect the fusion protein in transfected cells. A polyclonal antibody to the house-keeping protein glyceraldehyde 3-phosphate dehydrogenase (GAPDH) (37kDa) was used as a control. All cell lysates showed the presence of a 37 kDa band, corresponding to GAPDH, whereas only cells transfected with GFP fusion proteins showed a 64 kDa band, as shown in figure 6.4.

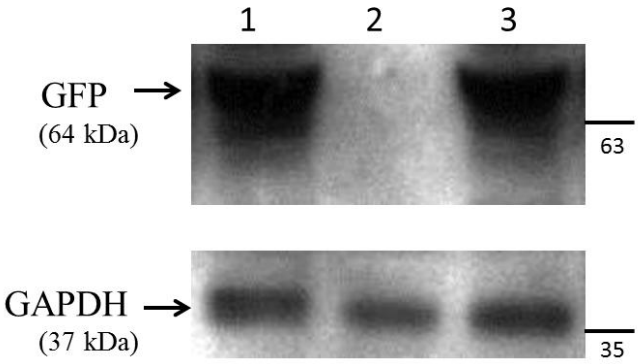


Figure 6.4. Western blot analysis of whole cell lysates from transfected cells

Western blots of whole cell lysates of HEK293T cells; 1) transfected with MC1R-wt, 2) untransfected and 3) transfected with MC1RΔ24, probed with GFP and glyceraldehyde 3-phosphate dehydrogenase (GAPDH) polyclonal antibodies. The GFP band (GFP fusion protein) corresponds to approximately 64 kDa, and the GAPDH corresponds to approximately 37 kDa, using 10-245 kDa protein ladder (Abcam). Horizontal lines indicate the positions of 63 and 35 kDa bands from the protein ladder.

6.2.4 Flow cytometry

Following transfection, cells were analysed using a flow cytometer (BD- FACSCalibur) to assess the number of cells expressing GFP. Figure 6.5 shows a typical histogram obtained

with cell counts plotted against fluorescence intensity. This number was used to measure the percentage of cells that were transfected and expressing the MC1R variant in each case. In a typical experiment, between 50%-60% of cells were transfected. Cells used for functional assays to determine intracellular cAMP levels were grown in parallel wells to those being assessed for transfection efficiency. Three wells were assessed by flow cytometry in this way for each MC1R variant in all experiments. The percentage of transfected cells in each case was used to normalise the data obtained from the functional assays.

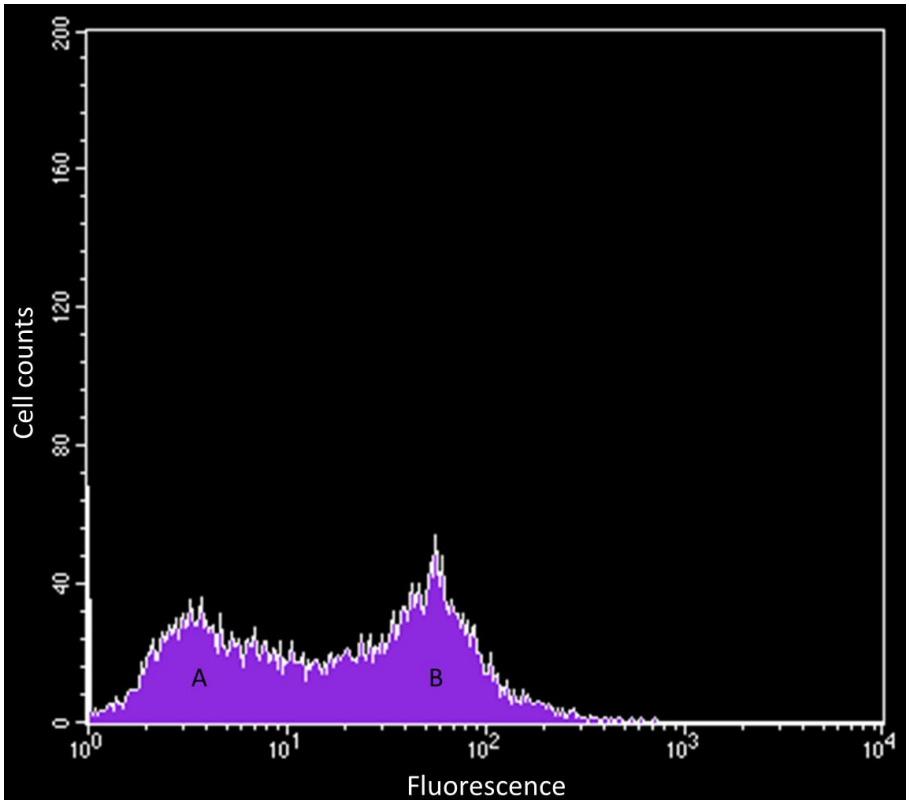


Figure 6.5. Histogram of transfected cells analysed by flow cytometry with cell count plotted against fluorescence

A representative histogram of transfected cells analysed by flow cytometry. Two distinct peaks show A) untransfected cells and B) transfected cells expressing GFP.

6.2.5 Functional assays to measure intracellular cAMP

6.2.5.1 Controls

All cAMP assays were carried out in triplicate and repeated in independent experiments (n=3). Controls were carried out for all experiments but, for clarity, these are not included in later graphs. Both untransfected cells and cells transfected with empty vectors were used as negative controls, and all cells were stimulated with 10 μM forskolin for 10 minutes as a positive control. Forskolin activates adenylate cyclase directly, raising cAMP levels, and as such is a control for the cell's ability to synthesise cAMP (Seamon, Padgett and Daly, 1981). Figure 6.6 shows the results of these controls, where untransfected cells and cells transfected with empty vectors produce levels of cAMP below 0.01 pmoles cAMP/μg protein and all cells produced approximately 2 pmoles cAMP/μg protein when stimulated with forskolin. All cAMP measurements presented in this study were normalised against transfection efficiency (where appropriate) and protein levels in cell lysates.

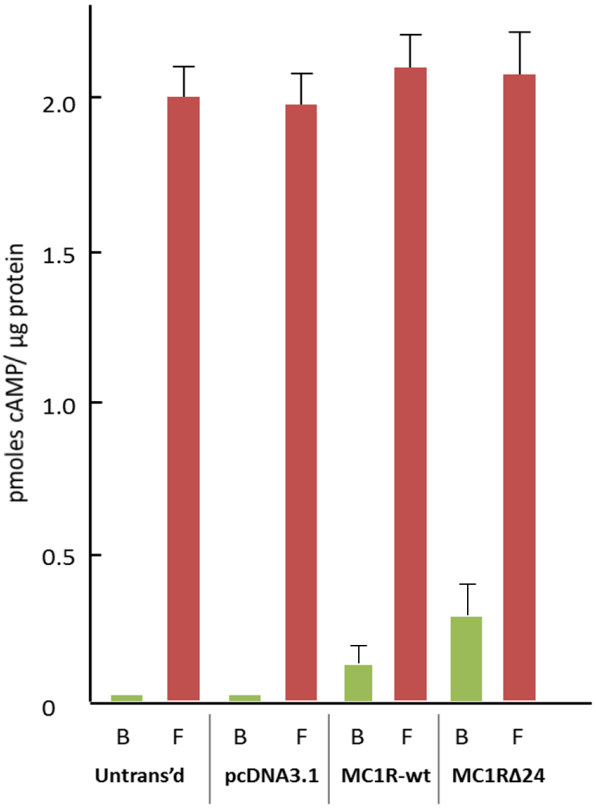


Figure 6.6. Functional coupling of melanocortin-1 receptor variants to intracellular cAMP levels in HEK293T cells stimulated with forskolin

Levels of intracellular cAMP from HEK293T cells when unstimulated (green bars), and when stimulated with 10µM forskolin (red bars). Untrans'd= untransfected cells, pcDNA3.1= cells transfected with empty vectors, MC1R-wt= cells transfected with vectors containing the MC1R-wt gene and MC1RΔ24= cells transfected with vectors containing the MC1RΔ24 gene, B= basal and F= forskolin. Error bars indicate standard error of the mean for n=3.

6.2.5.2 MC1R-wt and MC1RΔ24 response to α-MSH

Cells transfected with the *MC1R-wt* gene (MC1R-wt cells), showed a basal level of activity producing 0.1 pmoles cAMP/µg protein. This level represents 10% maximal response where

100% maximal response was the production of 0.95 pmoles cAMP/μg protein in the experiments from this study. The MC1R-wt cells showed a dose response to increasing concentrations of α-MSH, ranging from 0.1- 0.85 pmoles cAMP/μg protein (10% - 89% maximal response). The EC_{50} was 2.5×10^{-8} M, where EC_{50} is the concentration of the ligand which induces a half-maximal response. The MC1RΔ24 cells showed an elevated level of basal activity, producing 0.3 pmoles cAMP/μg protein (32% maximal response), as well as a dose response to α-MSH, ranging from 0.3- 0.94 pmoles cAMP/μg protein (32%- 99% maximal response) (see figure 6.7). The EC_{50} was 1.4×10^{-8} M.

6.2.5.2.1 Statistical analyses

Differences between the basal responses for the MC1R-wt and MC1RΔ24 cells were statistically significant at the 5% level (t (4) = -3.3, P = 0.03), but differences in maximal responses and EC_{50} were not statistically significant (t (4) = -0.89, P = 0.42 and t (4) = 1.3, P = 0.26 respectively).

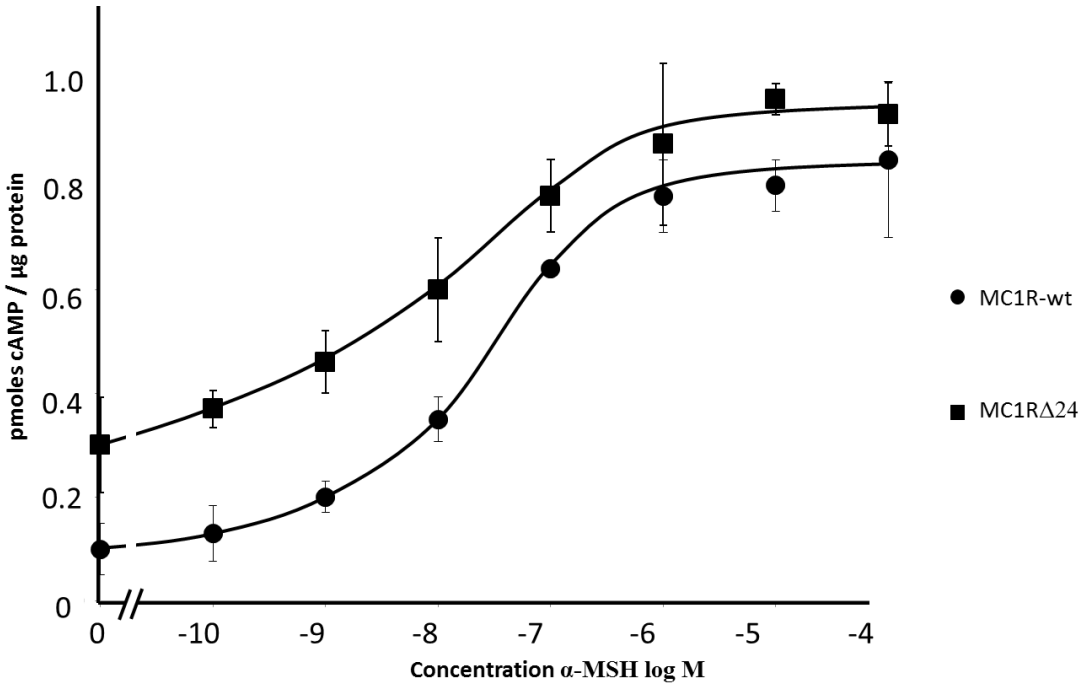


Figure 6.7. Functional coupling of MC1R-wt and MC1RΔ24 to intracellular cAMP levels in HEK293T cells stimulated with α-MSH

Changes to intracellular levels of cAMP in response to increasing concentrations of α-MSH in MC1R-wt or MC1RΔ24 transfected into HEK293T cells. Values on the y axis represent basal levels of activity in the absence of ligand and error bars indicate standard error of the mean for n=3. Differences between the basal responses for the MC1R-wt and MC1RΔ24 cells were statistically significant at the 5% level (t = -3.3, P = 0.03), but differences in maximal responses and EC50 were not statistically significant (t = -0.89, P = 0.42 and t = 1.3, P = 0.26 respectively).

6.2.5.3 MC1R-wt and MC1RΔ24 response to agouti signalling protein

The results of the response to α-MSH showed that the MC1RΔ24 receptor was constitutively active, and this alone could explain the black phenotype of squirrels with the *MC1RΔ24* allele, where elevated levels of cAMP lead to up-regulation of genes involved in eumelanogenesis as described in the introduction (Slominski et al., 2004). Although the MC1RΔ24 had an elevated level of basal activity, it was also still responsive to α-MSH, unlike some constitutively active receptors identified in mice (Robbins et al., 1993) and chickens (Ling et al., 2003), for example with the E92K mutation. The results of studies on the E92 mutant showed that basal levels are similar to maximal levels and receptors are no longer responsive to α-MSH. Here, in order to further investigate the effects of the deletion in the mutant, MC1RΔ24 transfected cells were stimulated with ASIP. The hypothesis was that the deletion in the MC1RΔ24 would lead to the receptor being unresponsive to its inverse agonist ASIP, whereas the wildtype receptor would respond in a dose-dependent manner.

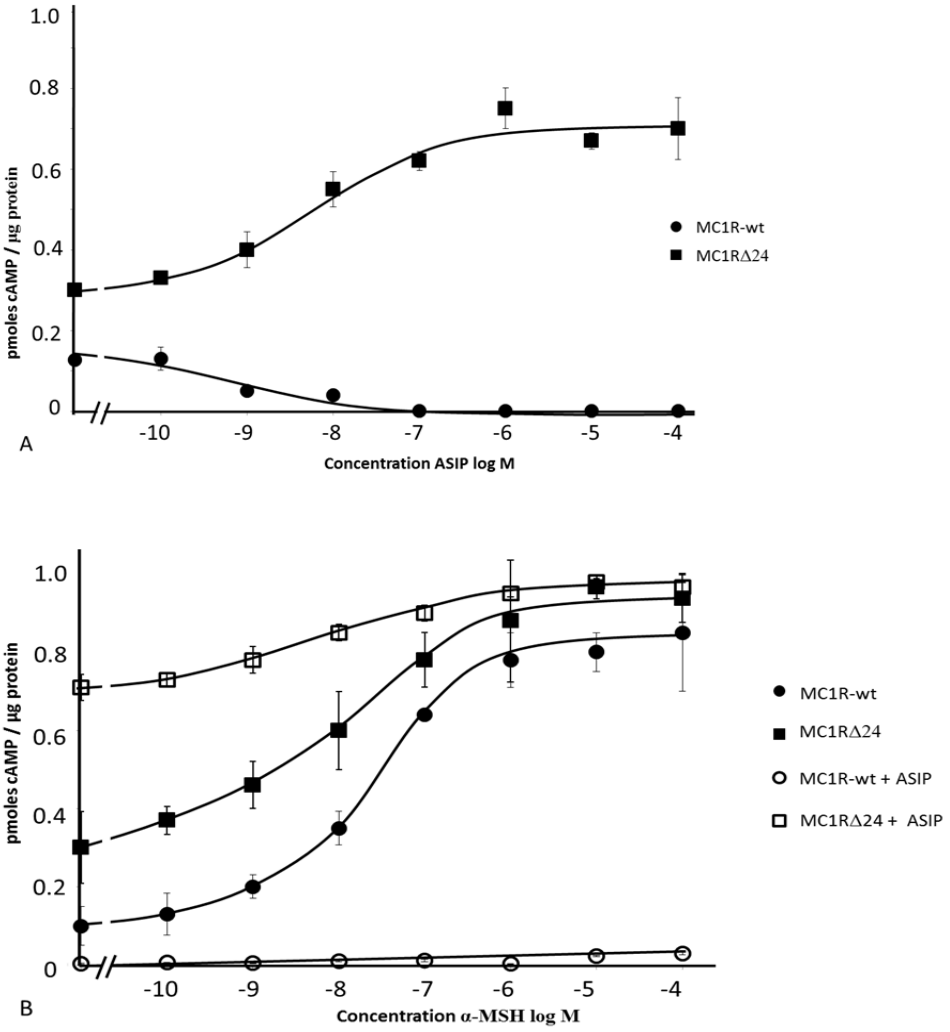


Figure 6.8. Functional coupling of MC1R-wt and MC1RΔ24 to intracellular cAMP levels in HEK293T cells stimulated with ASIP and/or α-MSH

A) Changes to intracellular levels of cAMP in response to increasing concentrations of ASIP in MC1R-wt and MC1RΔ24 transfected HEK293T cells. B) Changes in intracellular levels of cAMP in MC1R-wt or MC1RΔ24 transfected HEK293T cells, in response to increasing concentrations of α-MSH, with and without 100 000 nM ASIP. Values on the y axis represent basal levels of activity in the absence of ligand and error bars indicate standard error of the mean for n=3. Differences between maximal responses for MC1R-wt and MC1RΔ24 with ASIP were statistically significant at the 5% level (t = -4.32, P = 0.01). Differences between maximal responses for MC1R-wt and MC1RΔ24 cells with α-MSH and ASIP together were statistically significant at the 5% level (t =-15.52, P < 0.01). Differences between EC_{50} were not significant in either case (t = 0.74, P= 0.50 and t = -0.479, P = 0.66 respectively).

MC1R-wt cells showed a dose response to ASIP showing decreasing levels of cAMP production in response to rising concentration of ASIP, ranging from 0.13- 0.01 pmoles cAMP/μg protein (see figure 6.8). Contrary to expectations, MC1RΔ24 cells showed a positive dose response to ASIP with *increasing* cAMP production in response to rising concentration of ASIP, ranging from 0.3- 0.7 pmoles cAMP/μg protein (32%-74% maximal response) (see figure 6.8). These results show that α-MSH is an agonist for both MC1R variants, but that ASIP, whilst being an inverse agonist to MC1R-wt, is actually an agonist to MC1RΔ24. Further experiments, with saturating levels of ASIP (100 000 nM) and increasing concentrations of α-MSH, showed that the MC1R-wt cells exhibited an overall decrease in cAMP production, even at the highest concentration of α-MSH (100 000nM), with levels ranging from 0.03-0.04 pmoles cAMP/μg protein (3%- 4% maximal response) (see figure 6.8). In contrast, MC1RΔ24 cells showed an increased level of cAMP on stimulation with 100 000 nM ASIP, and cAMP production continued to rise with increasing concentrations of α-MSH, with levels ranging from 0.7- 0.95 pmoles cAMP/μg protein (74%-100% maximal response) (see figure 6.8). These results show that both ASIP and α-MSH act as agonists to the MC1RΔ24, where the highest levels of cAMP were produced when the cells were stimulated with both agonists.

6.2.5.3.1 Statistical analyses

Differences between maximal responses for MC1R-wt and MC1RΔ24 with ASIP were statistically significant at the 5% level (t (4) = -4.32, P = 0.01). Differences between maximal

responses for MC1R-wt and MC1RΔ24 cells with α-MSH and ASIP together were statistically significant at the 5% level (t (4) = -15.52, $P < 0.01$). Differences between EC_{50} were not significant in either case (t (4) = 0.74, P= 0.50 and t (4) = -0.48, P = 0.66 respectively).

6.2.5.4 MC1R-N88A response to α-MSH and agouti signalling protein

Computer models of the MC1R-wt and MC1RΔ24, predicted by Phyre2, suggest that the MC1R-wt has a hydrogen bond between N88 of TM2 and S124 of TM3, as discussed in chapter 5 (see figure 5.9). This hydrogen bond is missing in the MC1RΔ24, the N being one of the eight amino acids deleted in the mutant (SNALLETTI). The hypothesis here was that replacing the asparagine (N88) with an alanine (A88) would remove the hydrogen bonding with S124 and could create a more flexible receptor, and this in turn could lead to constitutive activity in the MC1R-wt (see figure 6.9). The functional assay was repeated with the MC1R-wt, MC1RΔ24 and the MC1R-N88A mutant.

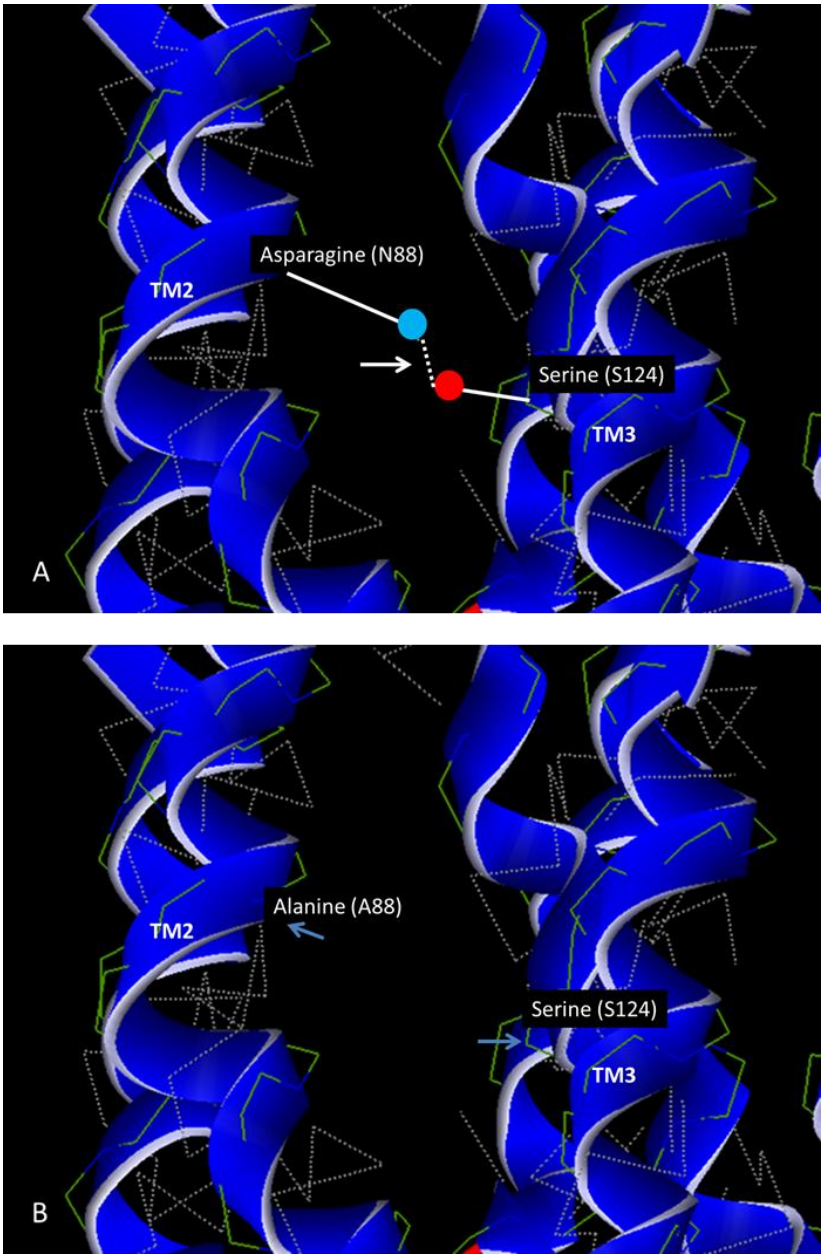


Figure 6.9. Computer models of the MC1R-wt and MC1R-N88A visualised with Hex 8.0

Close up of the hydrogen bonds between transmembrane helices two and three in the MC1R-wt. Hydrogen bonds are shown as dotted white lines. A) MC1R-wt with the hydrogen bond between N88 and S124 highlighted with a white arrow. Side chains of N88 and S124 are represented as white lines with nitrogen of the N88 as a blue circle and oxygen of the S124 as a red circle. B) Computer model of the MC1R-N88A mutant, showing A88 and S124 highlighted with blue arrows, and no hydrogen bond. TM = transmembrane helix.

The N88A mutation reduced the ability of the receptor to respond to α-MSH. Both the basal and maximal levels of cAMP were lower than those of the MC1R-wt with levels ranging from 0.07- 0.69 pmoles cAMP/μg protein (8% - 73% maximal response) but the receptor was still able to respond in a dose-dependent manner to increasing concentrations of α-MSH, as shown in figure 6.10.

6.2.5.4.1 Statistical analyses

Differences between basal responses for the MC1R-wt and MC1R-N88A cells were not statistically significant (t (4) = 0.97, P = 0.39). Differences in maximal responses to α-MSH were not statistically significant (t (4) = 1.82, P = 0.14) and differences for EC_{50} responses to α-MSH were significant (t (4) = -3.30, P = 0.03). Differences between maximal responses to ASIP were significant (t (4) = 3.67, P = 0.02) but EC_{50} responses to ASIP were not (t (4) = 1.01, P = 0.37).

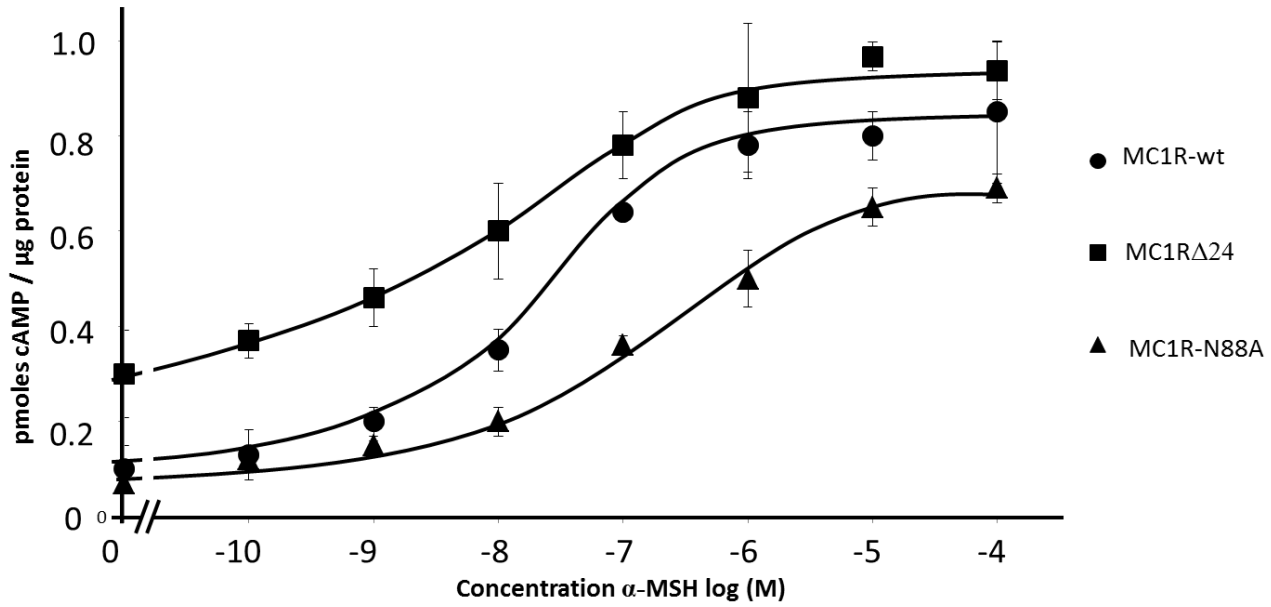


Figure 6.10. Functional coupling of MC1R-wt , MC1RΔ24 and MC1R-N88A to intracellular cAMP levels in HEK293T cells stimulated with α-MSH

Values on the y axis represent basal levels of activity in the absence of ligand and error bars indicate standard error of the mean for n=3. Differences between basal responses for the MC1R-wt and MC1R-N88A cells were not statistically significant (t = 0.97, P = 0.39). Differences in maximal responses to α-MSH were not significant t = 1.82, (P = 0.14) but differences for EC50 responses to α-MSH were (t = 3.67, P = 0.02).

The MC1R-N88A mutant showed a similar response to MC1R-wt with decreasing levels of cAMP production in response to rising concentrations of ASIP, with levels ranging from 0.06 - 0.001 pmoles cAMP/µg protein (1% - 0.2% maximal response) (figure 6.11).

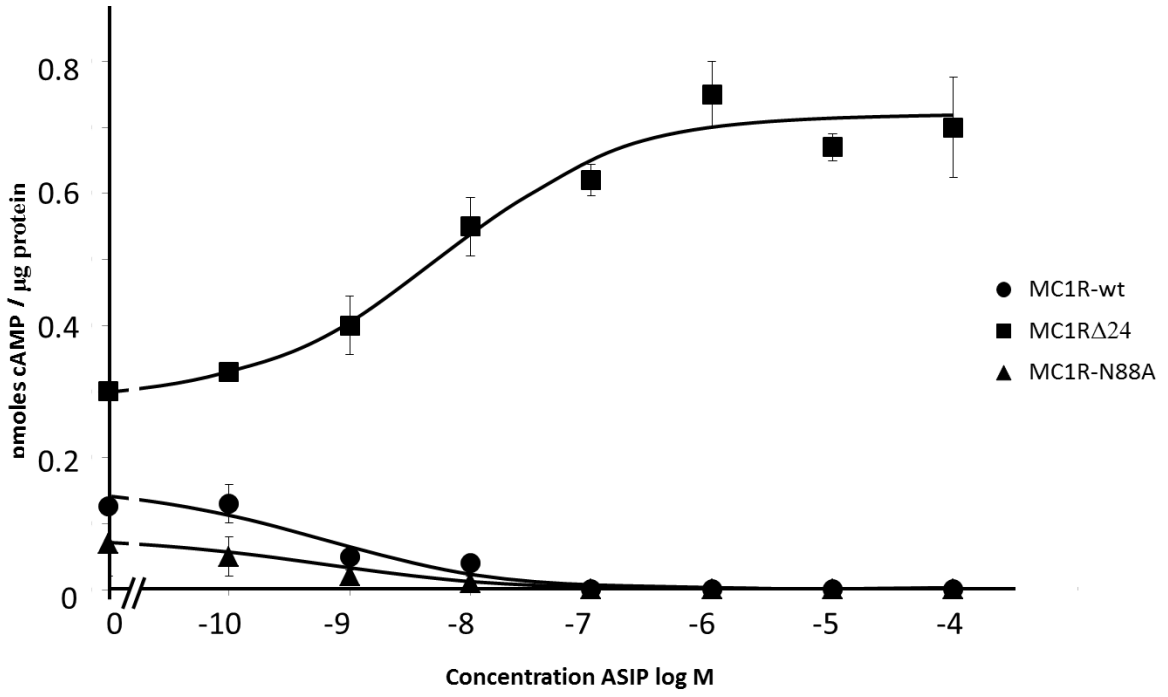


Figure 6.11. Functional coupling of MC1R-wt, MC1RΔ24 and MC1R-N88A to intracellular cAMP levels in HEK293T cells stimulated with ASIP

Values on the y axis represent basal levels of activity in the absence of ligand and error bars indicate standard error of the mean for n=3. Differences between maximal responses to ASIP for the MC1R-wt and MC1R-N88A cells were significant (t = 3.67, P = 0.02) but EC50 differences were not significant (t = 1.01, P = 0.37).

6.2.5.5 MC1R-del2 and MC1R-del4 responses to α-MSH and agouti signalling protein

To explore the effects of different deletions, two mutations were created with different deletions. Firstly, two amino acids were deleted; the SN of the SNALETTI sequence (87-94 MC1R-wt numbering) which is deleted in the MC1RΔ24. Secondly, four amino acids were deleted; the SNAL of SNALETTI. The structures of these mutants (MC1R-del2 and MC1R-del4 respectively) were predicted by the web server Phyre2 and these were used to create the schematic representations shown in figure 6.12.

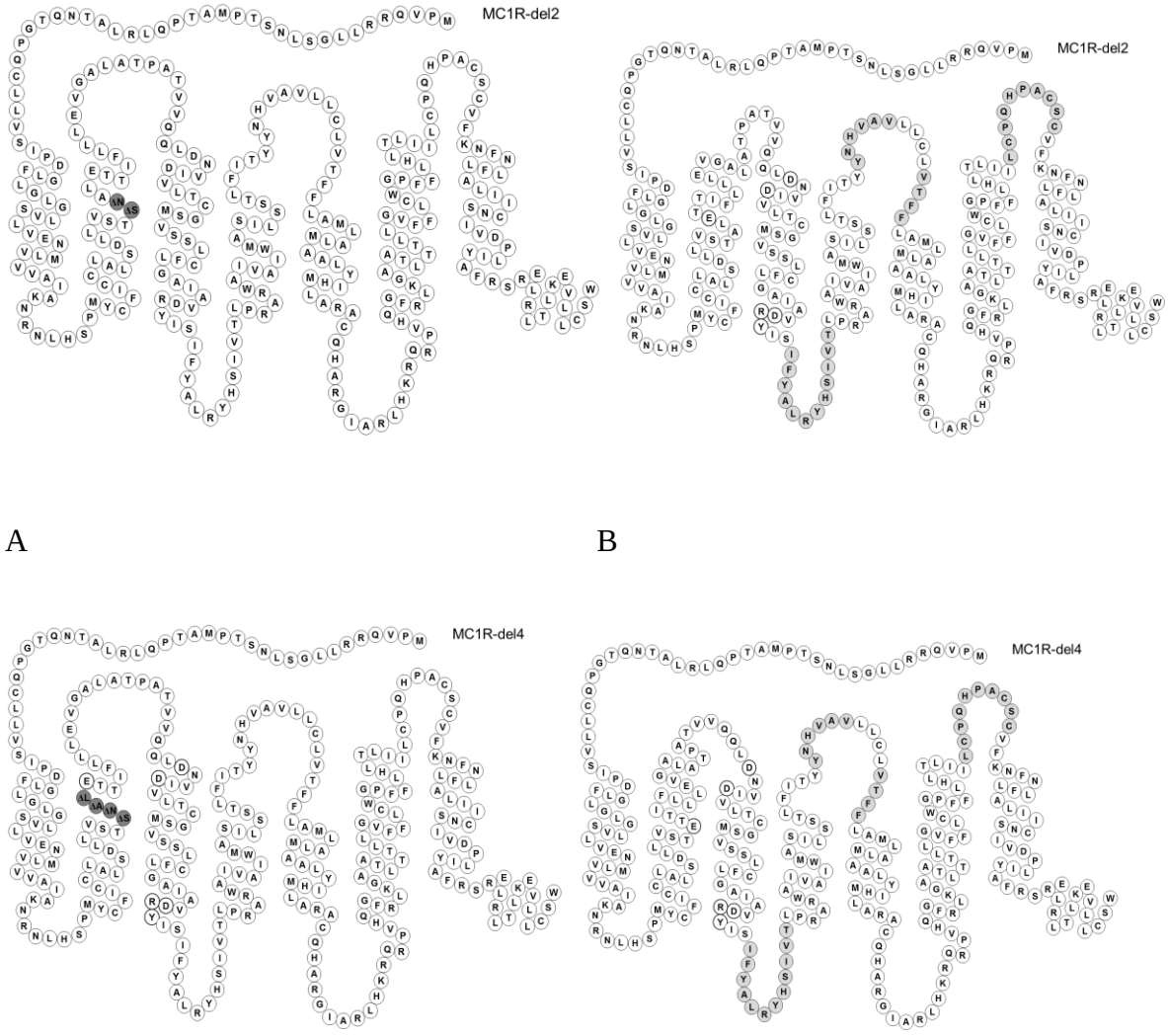


Figure 6.12. Schematic representations of MC1R-del2 and MC1R-del4

A) MC1R-wt showing the two amino acids to be deleted highlighted in grey, and B) showing the predicted structure of the MC1R-del2. C) MC1R-wt showing the four amino acids to be deleted, highlighted in grey, and D) showing predicted structure of the MC1R-del4. Predicted structures were obtained from the web server Phyre2.

The function of MC1R-del2 appeared to be completely abolished, and the receptor showed the same intracellular levels of cAMP as untransfected cells (< 0.01 pmoles cAMP/μg protein). The MC1R-del2 cells showed no response to α-MSH, as shown in figure 6.13.

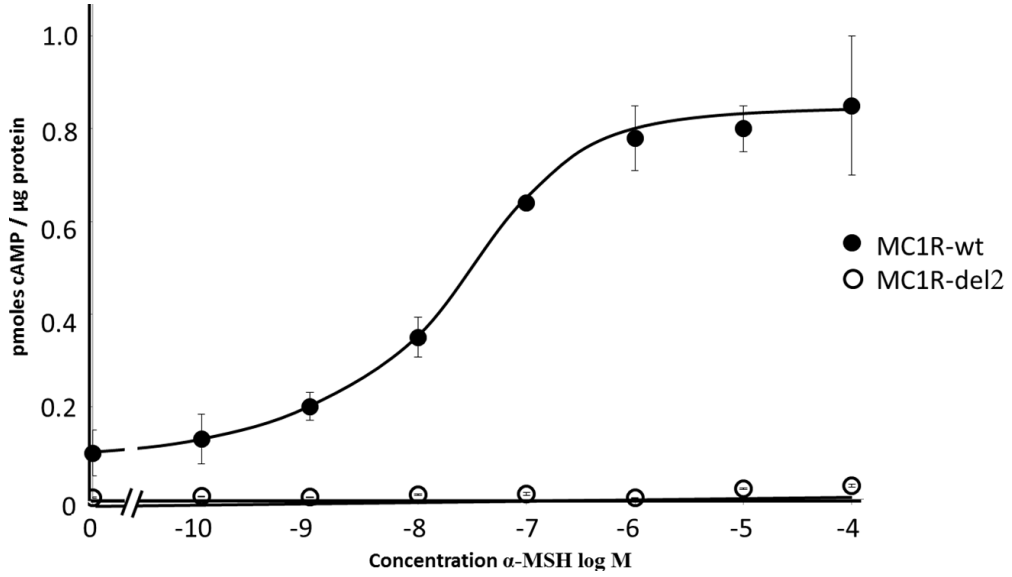


Figure 6.13. Functional coupling of MC1R-wt and MC1R-del2 to intracellular cAMP levels in HEK293T cells stimulated with α-MSH

Values on the y axis represent basal levels of activity in the absence of ligand and error bars indicate standard error of the mean for n=3. Differences between basal and maximal responses for the MC1R-wt and MC1R-del2 cells with α-MSH were both statistically significant (t = 3.4, P = 0.03 and t = 9.56, P = 0.03 respectively). EC_{50} differences were also statistically significant (t = -3.03, P = 0.04).

The four amino acids deleted in the MC1R-del4 mutant dramatically affected the receptor's function with an almost total loss of basal activity in MC1R-del4 cells, and a minimal response to α-MSH, with cAMP levels ranging from 0.02- 0.1 pmoles cAMP/μg protein (2% -11% maximal response), as shown in figure 6.14.

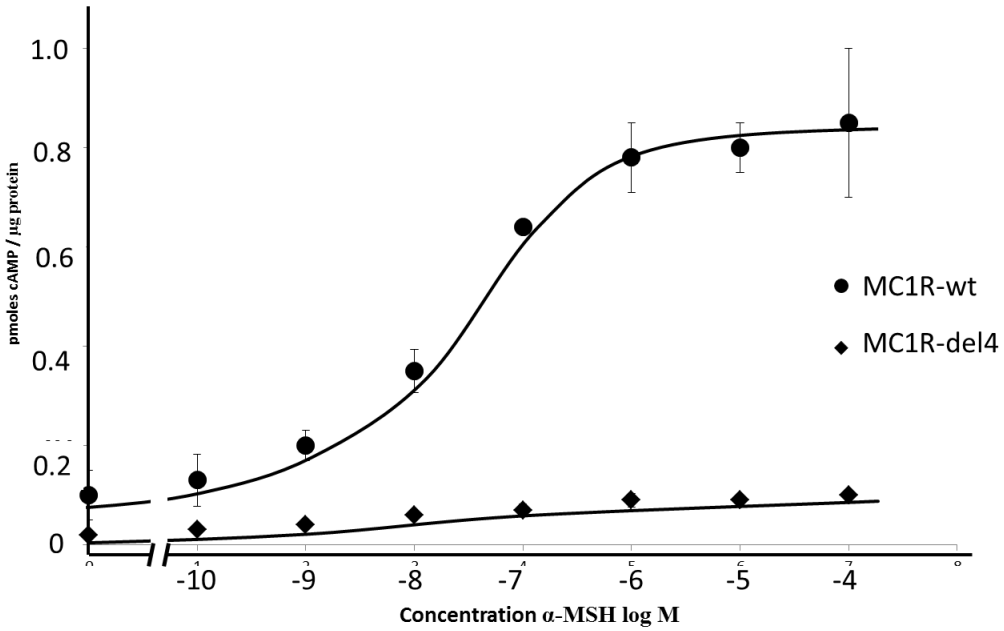


Figure 6.14. Functional coupling of MC1R-wt and MC1R-del4 to intracellular cAMP levels in HEK293T cells stimulated with α-MSH

Values on the y axis represent basal levels of activity in the absence of ligand and error bars indicate standard error of the mean for n=3. Differences between basal and maximal responses for the MC1R-wt and MC1R-del4 cells with α-MSH were both statistically significant (P < 0.01 in both cases). EC_{50} differences were also statistically significant (P < 0.01).

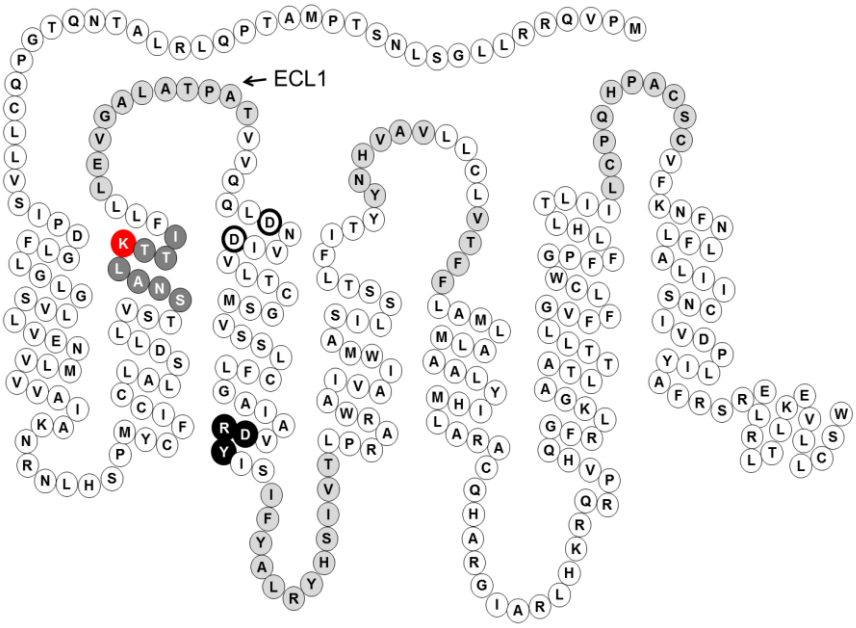
6.2.5.5.1 Statistical analyses

Differences between basal and maximal responses for the MC1R-wt and MC1R-del2 cells with α-MSH were both statistically significant (t (4) = 3.4, P = 0.03 and t (4) = 9.56, P = 0.03 respectively). Differences between basal and maximal responses for the MC1R-wt and MC1R-del4 cells with α-MSH were both statistically significant (t (4) = 2.87, P = 0.04 and t (4) = 8.71, P < 0.01 respectively). EC_{50} differences were statistically significant for MC1R-

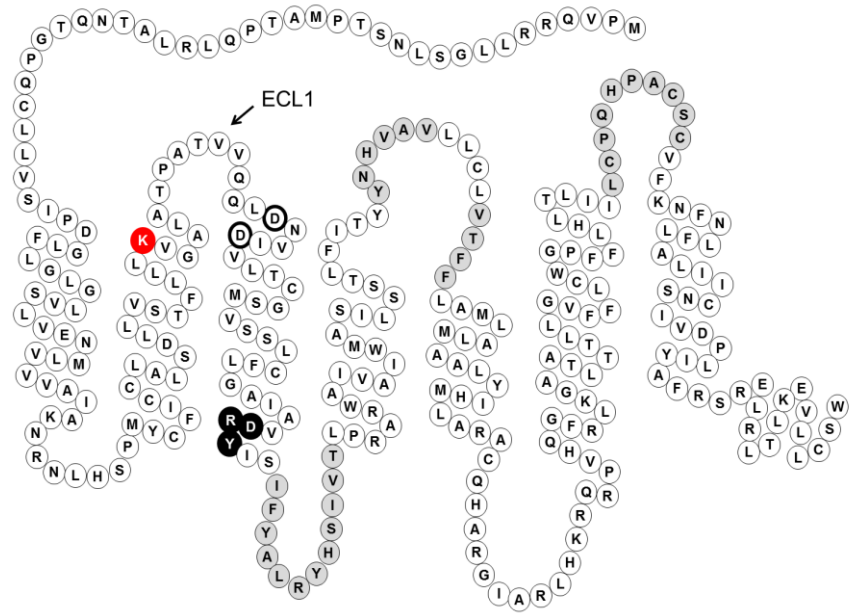
wt and MC1R-del2 cells with α-MSH but not for MC1R-wt and MC1R-del4 cells with α-MSH (t (4) = -3.03, P = 0.04 and t (4) = 0.19, P < 0.86 respectively).

6.2.5.6 MC1R-GE91Kand MC1R-BE91K responses to α-MSH and agouti signalling protein

To explore the function of the E91 residue of both the MC1R-wt and MC1RΔ24, two new mutants were created which changed the glutamic acid to a lysine, as shown in figure 6.15.



A



B

Figure 6.15. Schematic representations of MC1R-wt and MC1RΔ24 with the E91K mutation

A) MC1R-GE91K and B) MC1R-BE91K, showing the K91 as a red circle with white lettering. The DRY motif is highlighted as black circles with white lettering and the aspartic acids D114 and D118 (MC1R-wt numbering) are highlighted in bold.

Figure 6.15 A shows the MC1R-wt with the E91K mutation (MC1R-GE91K) and figure 6.15 B shows the MC1RΔ24 with the E91K mutation (MC1R-BE91K). The hypothesis was that changing the glutamic acid to a lysine would make both the MC1R-wt and MC1RΔ24 constitutively active as has been demonstrated for mice (Robbins et al., 1993; Benned-Jensen, Mokrosinski and Rosenkilde, 2011) and discussed above. Figure 6.16 shows that the MC1R-GE91K transfected cells were constitutively active, with a basal level of activity producing 0.38 pmoles cAMP/μg protein (40% maximal response). The MC1R-BE91K transfected cells had a reduced level of activity compared to the MC1RΔ24 cells, producing 0.14 pmoles cAMP/μg protein (15% maximal response) in all cases. Both E91K mutants were unresponsive to α-MSH or ASIP.

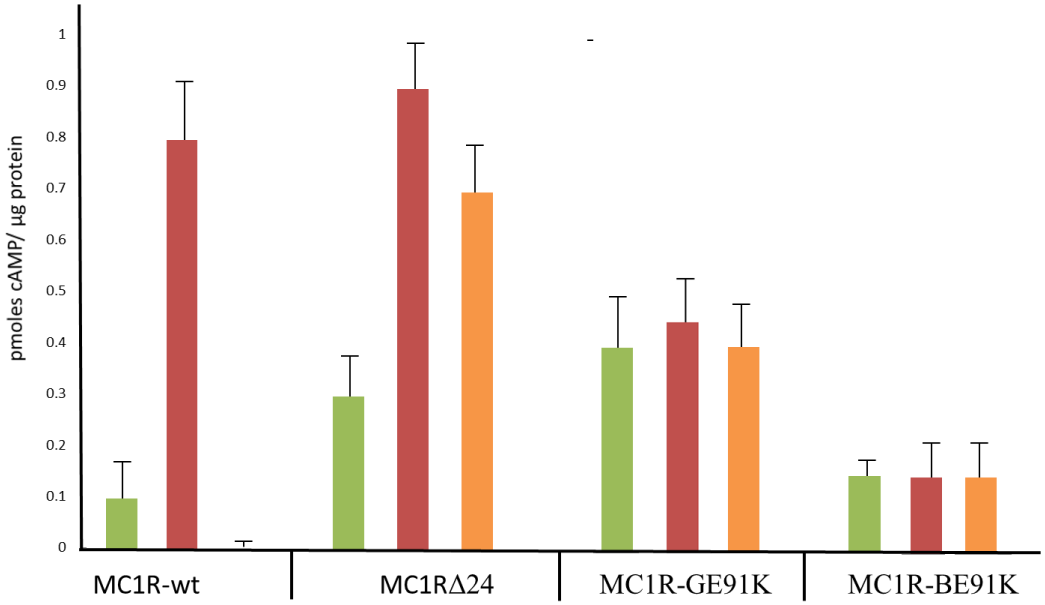


Figure 6.16. Functional coupling of MC1R E91K variants to intracellular cAMP levels in HEK293T cells stimulated with α-MSH and ASIP

The green bars indicate basal levels of cAMP, with no stimulation, red bars indicate cAMP levels of cells stimulated with 100 000 nM α-MSH and the orange bars indicate cAMP levels of cells stimulated with 100 000 nM ASIP. Error bars indicate standard error of the mean for n=3. Differences between the responses for the MC1R-wt, MC1RΔ24, MC1R-GE91K and MC1R-BE91K cells are summarised in table 6.1.

6.2.5.6.1 Statistical analyses

Differences between the responses for the MC1R-wt, MC1RΔ24, MC1R-GE91K and MC1R-BE91K cells are summarised in table 6.1.

Table 6.1. Summary of statistical t-test results comparing differences in intracellular cAMP levels for different E91K variants of the melanocortin-1 receptor

Basal refers to basal levels of cAMP, α-MSH max. refers to response at 100 000 nM α-MSH and ASIP max. refers to response at 100 000 nM ASIP.

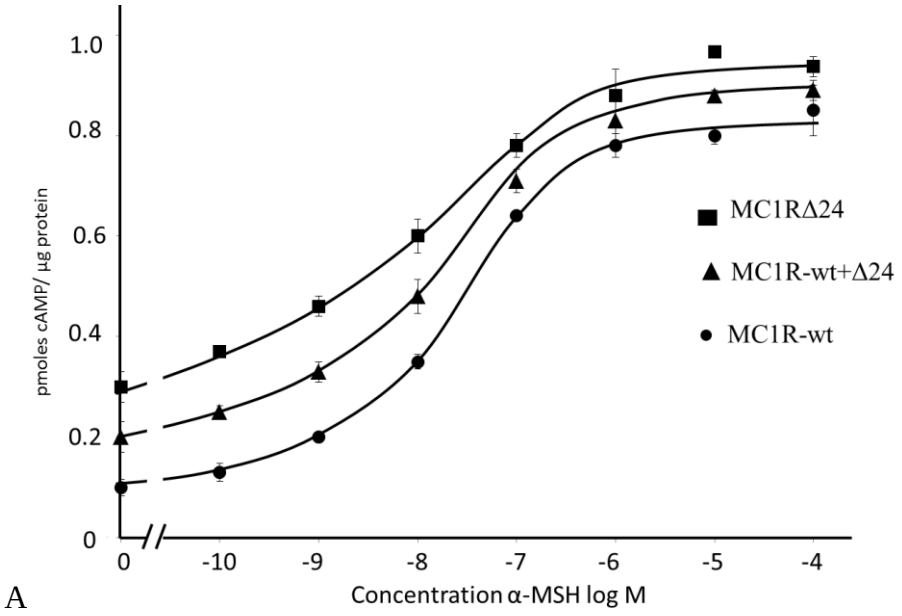
MC1R variants compared	**Conditions**	***t and P* values**	**Significant at 5%**
MC1R-wt and MC1R-GE91K	Basal	4.65, 0.03	yes
MC1R-wt and MC1R-GE91K	α-MSH max.	-4.32, 0.01	yes
MC1R-wt and MC1R-GE91K	ASIP max.	-4.36, 0.01	yes
MC1RΔ24 and MC1R-BE91K	Basal	-2.36, 0.06	no
MC1RΔ24 and MC1R-BE91K	α-MSH max.	16.20, <0.01	yes
MC1RΔ24 and MC1R-BE91K	ASIP max.	-3.90, 0.02	yes
MC1R-GE91K and MC1R-BE91K	Basal	3.90, 0.02	yes
MC1R-GE91K and MC1R-BE91K	α-MSH max.	4.00, 0.02	yes
MC1R-GE91K and MC1R-BE91K	ASIP max.	-3.35, 0.03	yes

6.2.5.7 Co-transfection of MC1R-wt and MC1RΔ24

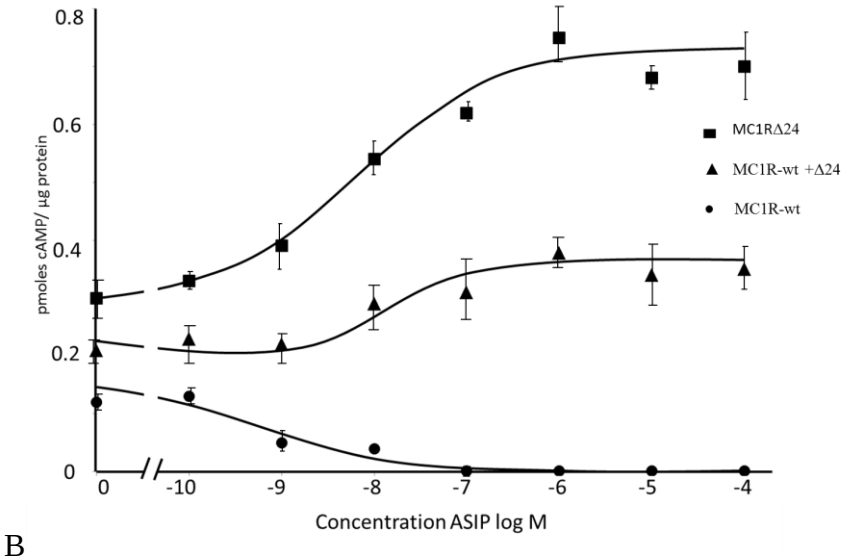
In order to simulate heterozygotes, (cells expressing both *MC1R-wt* and *MC1RΔ24* as found in the brown-black phenotype of the grey squirrel) cells were co-transfected with both MC1R-wt and MC1RΔ24 (MC1R-wt+Δ24). As expected, the results were intermediate, as shown in figure 6.17. MC1R-wt+Δ24 transfected cells had a basal level of activity producing 0.19 pmoles cAMP/μg protein (20% maximal response) and all other values were intermediate between the MC1R-wt and MC1RΔ24 cells.

6.2.5.7.1 Statistical analyses

Differences between basal responses for the MC1R-wt and the co-transfected MC1R-wt – MC1RΔ24 cells were not statistically significant (t (4) = -2.09, P = 0.11). Differences between maximal responses to α-MSH were not statistically significant (t (4) = -0.45, P = 0.68) but responses to ASIP were statistically significant (t (4) = -5.28, P = 0.01).



A



B

Figure 6.17. Functional coupling of MC1R-wt +Δ24 co-transfection to intracellular cAMP levels in HEK293T cells stimulated with α-MSH and ASIP

Changes in intracellular levels of cAMP in response to A) α-MSH and B) ASIP with MC1R-wt and MC1RΔ24 co-transfected into HEK293T cells. Values on the y axis represent basal levels of activity in the absence of ligand and error bars indicate standard error of the mean for n=3. Differences between basal responses for the MC1R-wt and the co-transfected MC1R-wt –MC1RΔ24 cells were not statistically significant (t = -2.09, P = 0.11). Differences between maximal responses to α-MSH were not statistically significant (t = -0.45, P = 0.68) but responses to ASIP were statistically significant (t = -5.28, P = 0.01).

6.3 Discussion

6.3.1 Constitutive activity of the MC1RΔ24

These results showed that the eight amino acid deletion in the MC1R leads to a constitutively active receptor producing higher basal levels of 0.3 pmoles cAMP/μg protein (32% maximal response) compared to the wildtype receptor which produced 0.1 pmoles cAMP/μg protein (10% maximal response). These levels of cAMP production are similar to comparable functional studies on human MC1R variants carried out using HEK293T cells, where basal levels of cAMP ranged from 0.05 – 0.15 pmoles cAMP/μg protein (Jiménez-Cervantes et al., 2001). The results from functional studies on the MC1R variants in the present study indicated that the equilibrium between R (inactive) and R* (active) had moved towards R* in the MC1RΔ24 receptor compared to the MC1R-wt receptor. The eight amino acid deletions were predicted to cause substantial structural changes to the receptor as described in chapter 5. The data from Socket predicted that both the MC1R-wt and the MC1RΔ24 had identical "knobs into holes" interactions (see chapter 5). This is perhaps not surprising as both receptors are functional and it has been shown that there is a consensus network of 24 inter-transmembrane helix contacts in both active and inactive GPCRs (Madabushi et al., 2004). It is therefore likely that these contacts are essential for GPCR function to be retained. It was hypothesised that the close association between TM2 and TM3 and ECL1 shortening in the MC1RΔ24 could lead to an upward shift of the TM2-TM3 bundle towards the extracellular side thus mimicking the upward movement observed in active receptors. This hypothesis is consistent with the finding that the MC1RΔ24 receptor had a higher basal activity than the wildtype receptor.

6.3.2 Activity of the MC1R-N88a mutant

Results for the MC1R-N88A mutant show that the receptor essentially behaves in the same way as the wildtype receptor but less effectively, having a lower basal activity and only reaching 60% maximal cAMP on maximum stimulation with α-MSH (figure 6.10). Contrary to predictions, this mutation did not make the receptor constitutively active but had a deleterious effect on function. The results suggest that the hydrogen bond between the N88 and S124 in the wildtype may be involved in, but not be essential to maintaining the function of the wildtype receptor. The loss of the bond may mean that movement of TM2 provoked during activation is not transmitted as effectively to TM3 and in turn not transmitted to the intracellular side of the receptor. The N88A mutation may have had other effects on the receptor, possibly leaving the S124 able to form other interactions. This result is consistent with the concept that TM2 and TM3 interactions are involved in activation of the MC1R.

6.3.3 Basal activity in MC1R-del2 and del4

Computer models of the MC1R-del2 and MC1R-del4 mutants show that these deletions could cause significant structural changes to the MC1R compared to both MC1R-wt and MC1RΔ24, as shown in figure 6.12 and 6.18 The model of MC1R-del2 predicts that the FLLLEVGAL (95-103 MC1R-wt numbering) sequence from the wildtype of ECL1 is incorporated into TM2 and that ECL1 is shortened to ATPATVVQ (104-111 MC1R-wt numbering) (figure 6.12). The model of MC1R-del4 predicts that the FLLLEVGALATPA (95-107 MC1R-wt numbering) sequence from the wildtype ECL1 is incorporated into TM2 and that ECL1 is shortened to TVVQQLD. The TM2s of MC1R-del2 and MC1R-del4 have

32 and 34 residues respectively, both of which are longer than those of MC1R-wt (25 residues) and MC1RΔ24 (27 residues) (figure 6.12).

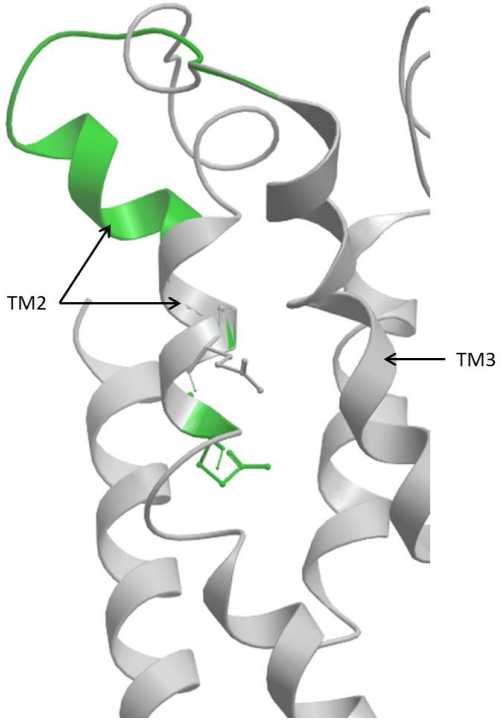


Figure 6.18. Detail from computer models of MC1R-wt and MC1R-del2 superimposed

Transmembrane helix two of the MC1R-del2 is predicted to be longer than that of the MC1R-wt, and the region of helix coloured green shows the lengthened region of MC1R-del2. The E91 residue is highlighted in each case as a stick model. This residue is relocated to position E89 in the MC1R-del2 and is coloured green while the E91 of the MC1R-wt is shown in grey. TM = transmembrane helix. Receptors are visualised with Molsoft.

The longer TM2 of MC1R-del2 and MC1R-del4 may create more intramolecular interactions and make the receptor more stable and less flexible than the wildtype or MC1RΔ24 receptors. This lengthening of TM2 in both cases would be consistent with the results which show that

these mutants both have no measurable basal activity. The lengthened TM2 could also be affecting ligand recognition and/or binding. These results showed that different lengths of deletions have different effects on helix formation and demonstrated that it is not the size of the deletion which is significant but the overall architecture of the receptor which affects function. These results demonstrated that a smaller deletion had a more profound and deleterious effect than a larger deletion. Examples of the effects of different sized deletions can be found in naturally occurring mutations; the results in the present study demonstrated that the 24 bp deletion of the squirrel caused a gain in function in the MC1R (McRobie et al., 2014b), but previous studies by Fontanesi et al. (2006) showed that a larger deletion of 30 bp deletion at the top of TM2 in the rabbit (MC1RΔ30) causes a loss-of-function in the receptor. Although there is no functional data for the rabbit MC1R, classical genetics show that this is a recessive mutation leading to red fur production in the homozygote (Fontanesi et al., 2006). The levels of cAMP from both MC1R-del2 and MC1R-del4 were comparable to untransfected cells and not likely to be high enough to support eumelanogenesis in melanocytes. Although the MC1R-del4 could respond to α-MSH, the concentration of agonist at which it responds is higher than any physiological levels. Interestingly, computer models predict that the position of the E91 residue (squirrel numbering) is the same for MC1R-wt and MC1RΔ24 from the squirrel and MC1RΔ30 from the rabbit (figure 6.19).

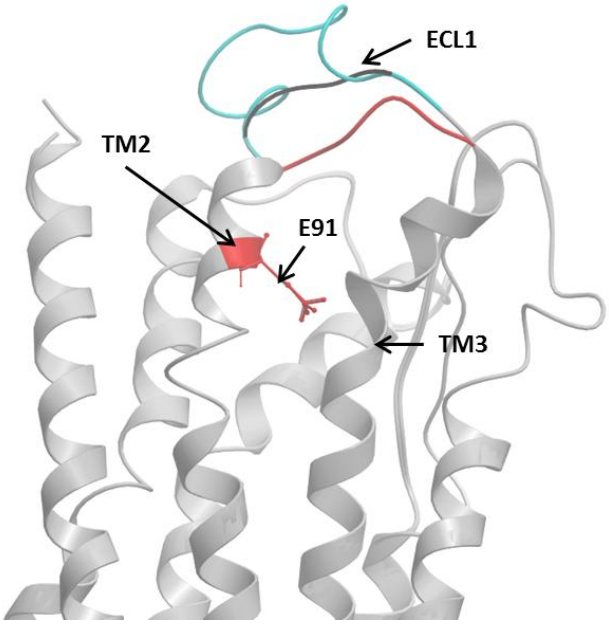


Figure 6.19. Detail of computer models of MC1R-wt, MC1RΔ24 and MC1RΔ30 superimposed

The extracellular loops one of each variant are highlighted in different colours: MC1R-wt in blue, MC1RΔ24 in black and MC1RΔ30 in red. E91 of all three receptors is shown in red as a stick model.

6.3.4 Receptor responses to α-MSH

The results presented in this study show that both MC1R-wt and MC1RΔ24 receptors respond to α-MSH with increased activity, in a dose-dependent manner. Physiological levels of α-MSH range from 0.01- 100 nM in humans (0.1 nM being the minimal effective dose), (Abdel-Malek et al., 1995), and it seems reasonable to expect a similar range to be found in squirrels. This would be consistent with the results presented here, where the greatest sensitivity to changes in α-MSH concentrations were between 0.1 and 100 nM, as shown in figure 6.20. The EC_{50} results showed that the MC1RΔ24 required less α-MSH to reach half maximal response than the MC1R-wt, but statistical tests showed that the differences were not significant between the two cases. Therefore it is concluded that α-MSH has similar effectiveness on both receptors.

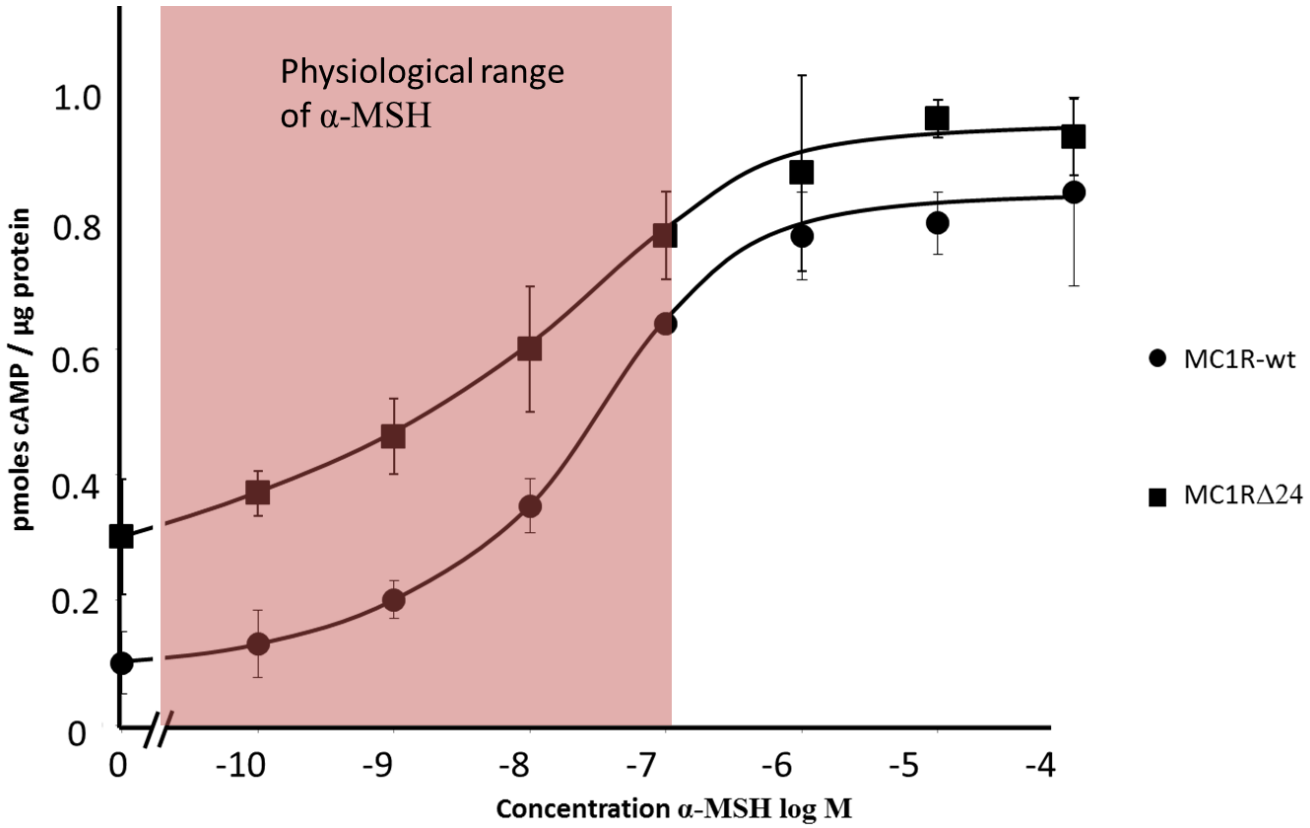


Figure 6.20. Functional coupling of MC1R-wt and MC1RΔ24 to intracellular cAMP levels in HEK293T cells showing the physiological range of concentrations of α-MSH

The physiological range of α-MSH in humans (0.01 -100 nM) is highlighted in pale red (Abdel-Malek et al., 1995).

Both MC1R-wt and MC1RΔ24 are predicted to have the same conserved residues in the binding pocket, as described in chapter 5. The residues E91, D114 and D118 (MC1R-wt numbering) which contribute to a highly charged, acidic region involved in ligand binding are found in both receptors but are likely to be slightly different in each case (Haskell-Luevano et al., 1996) figure 6.21. The finding that both wildtype and MC1RΔ24 receptors are responsive to α-MSH is consistent with both receptors having a similar acidic binding region. The ability

to respond to α-MSH is lost in the MC1R-del2 mutant and severely reduced in the MC1R-del4 mutant (figures 6.13 and 6.14). It appears that the MC1R-del4 receptor is stabilised in the inactive conformation but is able to respond to α-MSH at the highest concentrations with a modest response. Models predict a glutamic acid at the extracellular side of TM2 for both MC1R-del2 and MC1R-del4 but the position of this residue is deeper in the pocket and possibly inaccessible to ligands (figures 6.21 and 6.22).

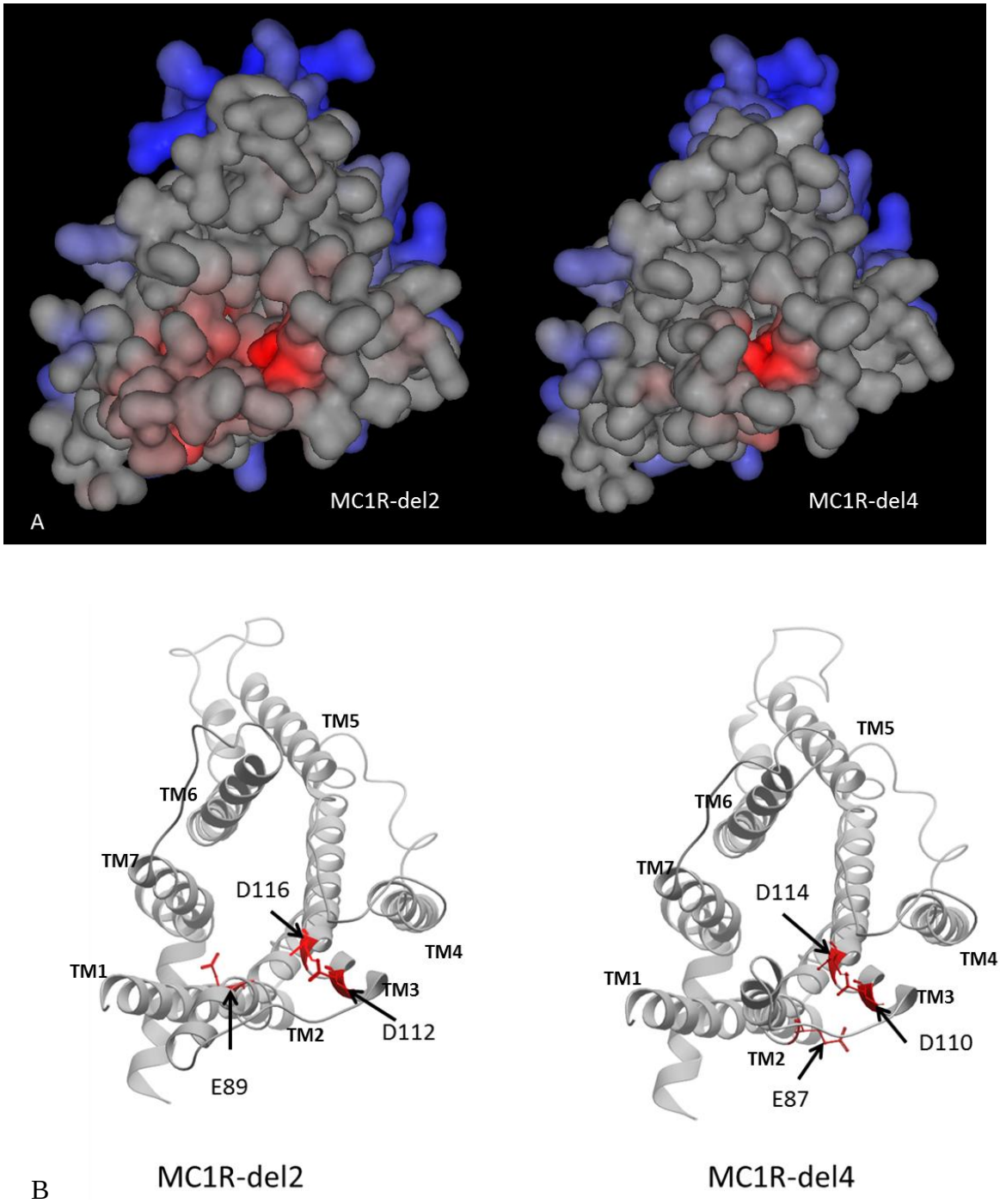


Figure 6.21. Computer models showing the extracellular view of MC1R-del2 and MC1R-del4

A) Molecular surfaces of the receptors with blue representing positive charges and red representing negative charges with intermediate values in grey. B) Ribbon diagram representations of the same orientation as A, showing the glutamic acids and aspartic acids highlighted in red as stick representations as predicted by Phyre2.

Figure 6.22 and 6.23 shows that the E89 (MC1R-del2) and E87 (MC1R-del4) (E91 equivalent) are orientated differently in each case.

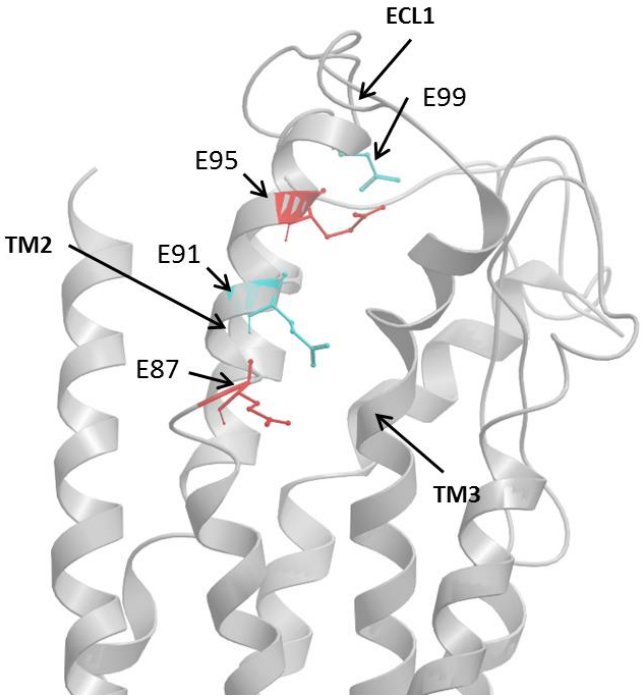


Figure 6.22. Detail of computer models of MC1R-wt and MC1R-del4 superimposed, showing the relative positions of the glutamic acids

Ribbon diagram representation of the MC1R variants showing stick representations of the E99 and E91 residues (coloured cyan) from the MC1R-wt and the equivalent E87 and E95 residues (coloured red) from the MC1R-del4 as predicted by Phyre2. TM = transmembrane helix, ECL = extracellular loop.

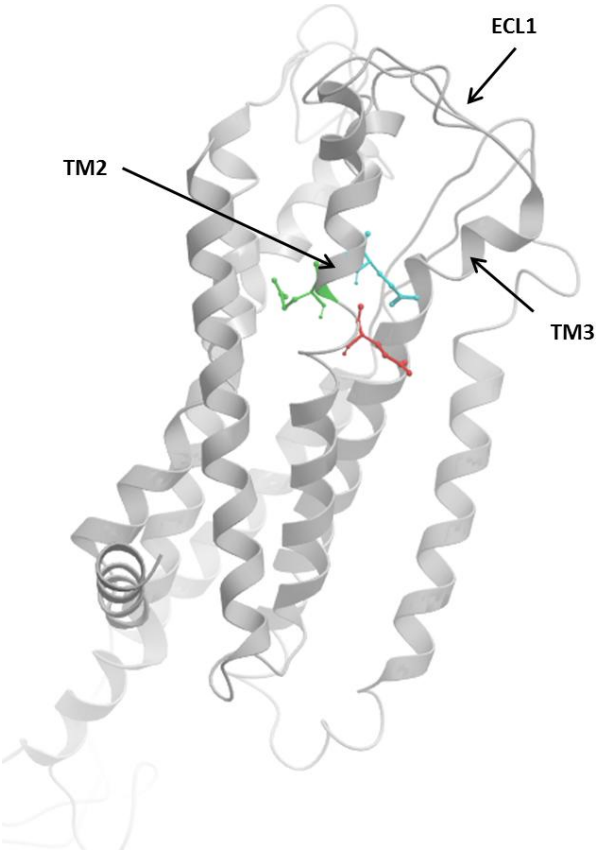


Figure 6.23. Computer models of MC1R-wt, MC1R-del2 and MC1R-del4 superimposed showing the relative positions of glutamic acids

Ribbon representation of the MC1R variants with the E91 from the MC1R-wt (blue), and the equivalent residues, E89 from MC1R-del2 (green) and E87 from MC1R-del4 (red) shown as stick representations as predicted by Phyre2.

The position and orientation of the residue in the MC1R-del2 case is significantly different to that of the wildtype, MC1RΔ24 and MC1R-del4. This change in orientation is likely to mean that the receptor can no longer be activated by α-MSH, as the acidic pocket is no longer maintained and other interactions are likely to be occurring (figure 6.23). The orientation of the E87 in the MC1R-del4 is not as dramatically changed compared to the MC1R-del2 and it

may be hypothesised that at high concentrations of α-MSH there could be an induced fit allowing the receptor to be stabilised in the R* conformation.

Figure 6.24 shows that many different MC1R receptors from a large range of species retain the E91 (squirrel numbering) in the same position. All of the receptors superimposed in the figure are functional receptors. These results highlight the importance of the acidic binding pocket in responding to α-MSH and in particular the importance of the E91 residue (Lu, Vage and Cone, 1998).

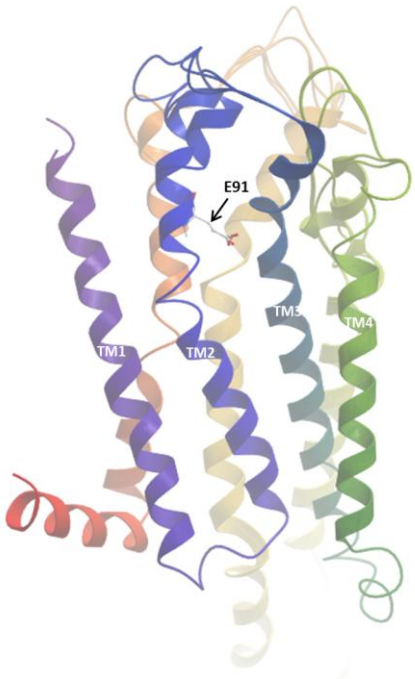


Figure 6.24. Computer model of the MC1R from the golden-headed lion tamarin, jaguar, jaguarundi, falcon, and squirrel superimposed on each other

The E91 (squirrel numbering) is highlighted as a stick representation in each case. Models were obtained from the web server Phyre2. Sequence information was obtained from Genbank (Eizirik et al., 2003; Mundy and Kelly, 2003; Gangoso et al., 2011).

6.3.5 The effect of the E91K mutations

The importance of the E91 residue is further emphasised by the results of the E91K mutants (figure 6.16). As expected from previous studies on other species, the E91K mutation in the wildtype receptor results in a constitutively active receptor which is unresponsive to both α-MSH and ASIP. Contrary to expectations the E91K mutation in the MC1RΔ24 receptor reduced basal activity. This receptor was also unresponsive to ligands and so it seems that the introduction of a positive charge may prevent ligands from entering the binding pocket. Figure 6.25 shows that the K91 is differently orientated between the two cases (as predicted by Phyre2) which may explain the different activities.

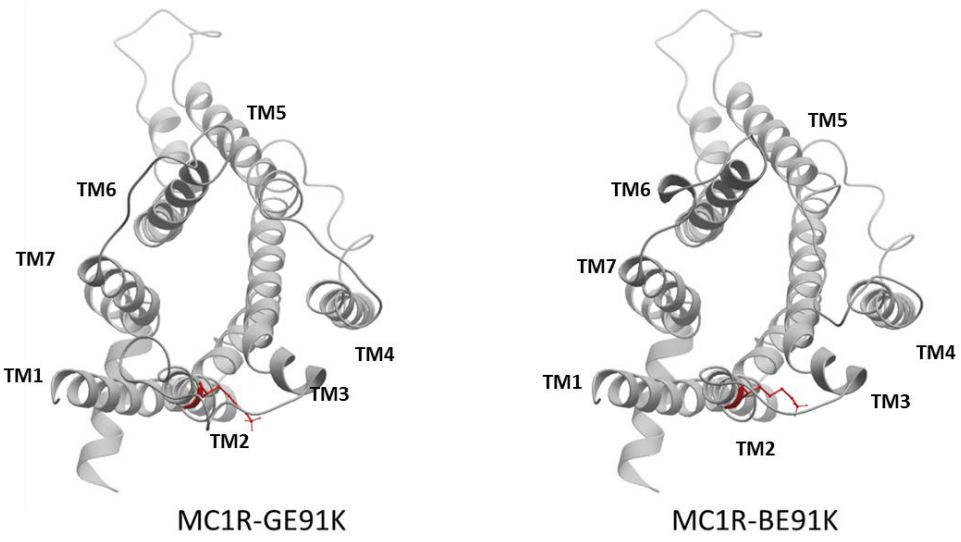


Figure 6.25. Computer models showing the extracellular view of the MC1R-GE91K and MC1R-BE91K with the K91 residue highlighted

Ribbon diagram representations of the MC1R variants with the K91 residue highlighted in red as a stick representation as predicted by Phyre2.

A number of species have the E92K (mouse numbering) mutation; for example mice, chicken, Japanese quail, the bananaquit and ruffed lemur (*Varecia variegata*) (Robbins et al., 1993; Takeuchi et al., 1996; Theron et al., 2001; Mundy and Kelly, 2003; Nadeau, Minvielle and Mundy, 2006). In the mouse, chicken and lemur, this mutation is known to lead to a constitutively active receptor (Robbins et al., 1993; Ling et al., 2003; Haitina et al., 2007). Mutagenesis studies on the E92 position confirm that a change from an acidic to a basic amino acid, either Lys or Arg, leads to constitutive activity. Binding studies show however, that the E92K reduces ligand binding (Lu, Vage and Cone, 1998; Haitina et al., 2007). It therefore seems likely that E92, D117 and D121 (mouse numbering) are in close proximity**,** but that they contribute to electrostatic repulsion between TM2 and TM3. The introduction of a basic residue could reduce this repulsion and so alter the overall structure of the receptor**,** which might then lead to receptor activation. Indeed, in the ruffed lemur, the E92, D117 and D121 residues are changed to K92, G117 and N121 (mouse numbering). Here again it seems that a loss of electrostatic repulsion, caused by the loss of positively charged residues at these positions, may be leading to the constitutive activation observed (Mundy and Kelly, 2003; Haitina et al., 2007). A similar situation may be occurring with the Goeldi's monkey (*Callimico goeldii*), where the MC1R is also constitutively active and where the D117 is changed to H117, introducing a positively charged His into the acidic pocket (Mundy and Kelly, 2003; Haitina et al., 2007). Similarly**,** the C125R mutation associated with constitutive activation and melanism in the fox would introduce a positive charge and is likely to have a similar agonist mimicking effect (Våge et al., 1997b). The L98P and L99P mutations associated with constitutive activity and melanism in mice and pigs, which introduce a proline, would likely modify the TM positions and again the effect would be transmitted along the TM3 (Robbins et al., 1993; Kijas et al., 1998). Thus**,** it seems likely that the effects

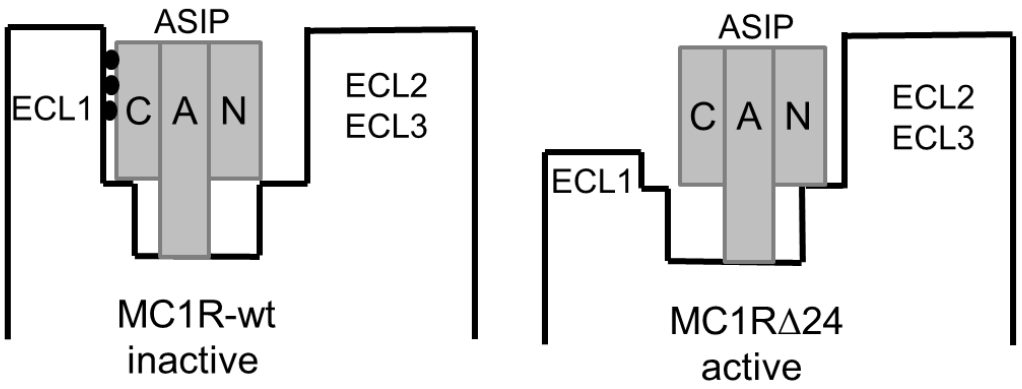
of the E92K, C125R, L98P and L99P mutations are mimicking the action of the Arg residue of α-MSH (Lu, Vage and Cone, 1998). Studies on the E92K mutation show that the active conformation induced by ligand binding is distinct from that induced by the mutation and that the extracellular loop one (ECL1) is involved in maintaining the active conformation (Benned-Jensen, Mokrosinski and Rosenkilde, 2011). The effects of these mutations may transmit a conformational change along the TM2/TM3 bundle and ultimately can be envisaged to transmit a structural signal to the intracellular side of the receptor and in this way affect G-protein specificity, turnover and/or binding energies.

The relocation of E99 to E91 (MC1R-wt numbering) in the MC1RΔ24 is likely lead to different interactions in the mutant. It is likely that the E91 will have different interactions and this may lead to an altered repulsion between TM2 and TM3 which could affect activity. The deleted amino acids may have the effect of creating a more stable receptor with an altered architecture and a permanently enlarged binding region for the G-protein so that it is more often in the active R* state. As previously mentioned, G-protein coupling on the intracellular side of GPCRs consistently increases the affinity of agonists and full activation requires both agonist and G-protein engagement (Audet and Bouvier, 2012). This may be relevant in the MC1RΔ24 where a higher basal activity may increase the likelihood of agonist binding. The alteration in the overall structure of the receptor could be leading to biased signalling. Future experiments measuring β-arrestin recruitment could establish whether the active conformations of MC1R-wt and MC1RΔ24 have similar characteristics or whether active conformations of each lead to different downstream signalling. Overall these findings highlight how the MC1R is a carefully balanced switch whose activity can easily be increased

and decreased with small changes. The changed position of E91 could also contribute to an altered interaction with ASIP.

6.3.6 Agouti signalling protein as agonist

The most unexpected finding in this study was that ASIP acts as an agonist to the MC1RΔ24 receptor. Given the finding that ECL1 is vital for ASIP to act as an inverse agonist and that ASIP with no C-terminus loop acts as an agonist, it seems likely that the shorter ECL1 of MC1RΔ24 is preventing ASIP from inactivating the receptor and instead the ASIP is able to bind to the same region as the α-MSH and activate the receptor (Figures 5.27 and 6.26).



A B

Figure 6.26. Schematic representation of ASIP-MC1R-wt and ASIP-MC1RΔ24 interactions

A) Predicted hydrophobic interactions between the C-terminal loop of ASIP and extracellular loop 1 of MC1R-wt are shown as black circles. B) Predicted relocation of ASIP when interacting with MC1RΔ24, leading to activation of the receptor, (adapted from Patel et al. (2010). ECL = extracellular loop, C= C-terminal loop, A =Active loop and N = N-terminal loop of ASIP. Figure taken from McRobie et al. (2014b).

Interestingly, the melanic jaguar, jaguarundi, and falcon have deletions in this region of the MC1R (table 6.2) (Eizirik et al., 2003; Gangoso et al., 2011). Additionally the golden-headed lion tamarin which is melanic except for the head also has a deletion in this region (Mundy and Kelly, 2003).

Table 6.2. Alignment of amino acids of melanocortin-1 receptor

Amino acid alignments of MC1R sequences (80-112) from the wildtype (accession number EU604830.3) and melanic (accession number EU604831.2) squirrel, wildtype (accession number AY237396) and melanic (accession number AY237397) jaguar and wildtype (accession number AY237399) and melanic (accession number AY237398) jaguarundi (Eizirik et al., 2003). Numbering is according to squirrel MC1R-wt. Deleted amino acids are shown as black rectangles. Dashes represent consensus with the wildtype squirrel sequence. Coloured blocks indicate amino acids predicted by computer modelling to be in extracellular loop 1.Table taken from McRobie et al. (2014b).

Species and phenotype	Amino acid sequence alignment																													
	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112
Squirrel-wt	S	D	L	L	V	S	T	S	N	A	L	E	T	T	I	F	L	L	L	E	V	G	A	L	A	T	P	A	T	V	V	Q	Q
Squirrel-mel	-	-	-	-	-	-	-	■	■	■	■	■	■	■	■	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Jaguar-wt	-	-	-	-	-	-	V	-	S	V	-	-	-	A	V	M	-	-	-	-	A	-	T	-	-	G	R	-	A	-	-	-	-
Jaguar-mel	-	-	-	-	-	-	V	-	S	V	-	-	-	A	V	M	-	-	T	■	■	■	■	■	-	G	R	-	A	-	-	-	-
Jaguarundis-wt	-	-	-	-	-	-	V	-	S	V	-	-	-	A	V	M	-	-	-	-	A	-	T	-	-	G	R	-	A	-	-	-	-
Jaguarundis-mel	-	-	-	-	-	-	V	-	S	V	-	■	■	■	■	■	■	■	■	-	A	-	T	-	-	G	R	-	A	-	-	-	-

Although these deletions fall in slightly different regions, computer models predict that the effects on ECL1 are similar (figures 6.24 and 6.27). Indeed, in all five species, ECL1 contains a section of β-strand and hydrophobic residues important for inverse agonist activity in the wildtype. Figure 6.27 shows that all ECL1s are shortened in the melanic phenotypes with no β-strand and fewer hydrophobic residues. It seems likely that these deletions are having similar effects to those found in the squirrel, causing changes to agonist and inverse agonist activity and causing constitutively active receptors. Further mutational studies on the MC1R and/or ASIP may determine the nature of the interactions between these molecules and reveal which residues are involved when ASIP acts as an inverse agonist and which interactions occur when it acts as an agonist.

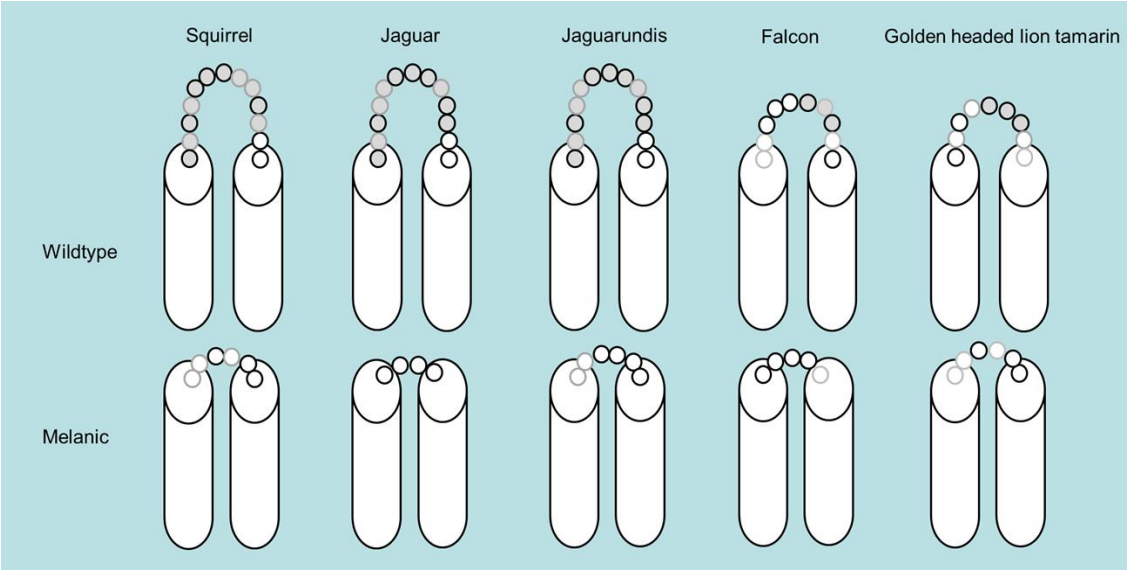


Figure 6.27. Schematic representation of TM2, TM3 and ECL1 from the MC1R of wildtype and melanic squirrel, jaguar, jaguarundi, falcon and golden-headed lion tamarin

Amino acids that form part of β-pleated sheets are shown as grey circles. Hydrophobic amino acids are shown as circles with black outlines.

6.3.7 Heterozygotes with MC1R-wt and MC1RΔ24

Cells co-transfected with MC1R-wt and MC1RΔ24 produced intermediate results to those for cells transfected with MC1R-wt and MC1RΔ24 when stimulated with α-MSH and ASIP (figure 6.17). This is not surprising as the individual cells are likely to express either MC1R-wt or MC1RΔ24 and not both and therefore the results are simply an average of the two cell types. However, this result does not give an accurate reflection on what could be occurring in heterozygous melanocytes *in vivo*. *In vivo*, the two receptor types would both be expressed in the same melanocyte and this could lead to interactions which cannot be observed in co-transfected cells. These results are, however, consistent with the phenotypic analysis of hairs. It is anticipated that the pulse of ASIP expression which occurs during the hair growth cycle

in mice (Vrieling et al., 1994), also occurs in squirrels, accounting for the banding pattern observed in squirrel hairs. These results predict that the MC1R-wt in melanocytes would respond to α-MSH and ASIP producing varying levels of intracellular cAMP, ultimately leading to the production of banded phaeomelanin/eumelanin hairs of the wildtype grey squirrel. In contrast, melanocytes with MC1RΔ24 may only be capable of producing eumelanin leading to the production of unbanded jet black hairs of the melanic grey squirrel. These results may account for the shorter phaeomelanin band of the heterozygote (figure 3.3) as melanocytes expressing both MC1R-wt and MC1RΔ24 would likely have an intermediate level of intracellular cAMP. Given that cAMP levels for MC1RΔ24 cells are always higher than MC1R-wt cells, heterozygotes are likely to always have higher cAMP levels than MC1R-wt cells. As a consequence, heterozygous melanocytes would reach the threshold for switching from eu- to phaeomelanogenesis more slowly than wildtype melanocytes during a pulse of ASIP expression, resulting in a band starting later in the hair growth cycle, as shown in figure 6.28. Similarly, eumelanogenesis would resume more quickly when the pulse has finished, ultimately leading to a shorter band of phaeomelanin compared to the wildtype. Figure 6.28 shows a representation of how a shorter band of phaeomelanin would be produced by heterozygotes. Point A on figure 6.28 shows the point at which the homozygous wildtype melanocytes would switch from eu-to phaeomelanogenesis and point D when eumelanogenesis would resume, producing a long band of orange pigment. Point B on the graph shows the point at which the heterozygous melanocyte would switch from eu- to phaeomelanogenesis and point C when eumelanogenesis would resume, producing a short band of orange. The homozygous MC1RΔ24, in contrast, does not cross the threshold resulting in solid black hairs.

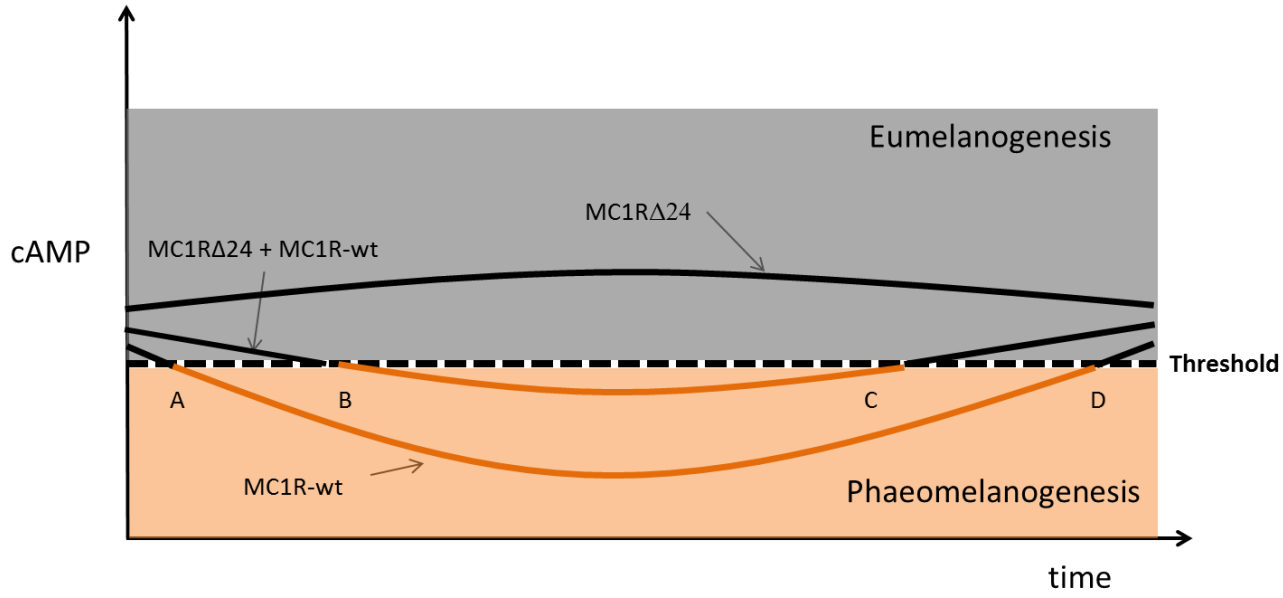


Figure 6.28. Schematic graph to show levels of cAMP during a pulse of agouti signalling protein expression for MC1R-wt, MC1RΔ24 and heterozygotes

The broken black line indicates the threshold level of cAMP for switching from eumelanogenesis to phaeomelanogenesis. Point A shows the time at which the MC1R-wt switches from eu-to phaeomelanogenesis and point D where eumelanogenesis is resumed. Point B shows the time at which the heterozygote switches from eu-to phaeomelanogenesis and point C where eumelanogenesis is resumed.

The ventral regions of the different colour morphs are particularly revealing as the wildtype is white, the heterozygote is orange and the homozygote jet black squirrel has a completely black ventrum (figure 3.2). This region of the body expresses ASIP continuously in other species and it could be predicted that the same is the case in squirrels. Assuming this is the case it could be predicted that melanocytes of the ventral in wildtype squirrels would have low levels of cAMP, leading to very low levels of melanogenesis, or complete cessation of

melanin production. In contrast, the ventral area of the jet black squirrel is predicted to have high levels of cAMP leading to eumelanogenesis whereas the heterozygote would have intermediate levels of cAMP leading to phaeomelanogenesis. Further studies on squirrel melanocytes and *in vivo* studies would be needed to confirm these predictions.

6.4 Conclusion

The eight amino acid deletion in the MC1RΔ24 leads to a constitutively active receptor, producing a higher basal level of intracellular cAMP compared to the wildtype receptor. α-MSH acts as an agonist to the MC1R-wt and MC1RΔ24 receptors while ASIP acts as an inverse agonist to the MC1R-wt and as an agonist to the MC1RΔ24 receptor. These results highlight the critical role of the glutamic acid at position 91 (MC1R-wt numbering).

Chapter Seven

General Discussion

7.1 Summary of research findings

Three distinct phenotypes of the grey squirrel were identified; the wildtype grey, the brown-black and the jet black (chapter 3). Melanism in the grey squirrel was found to be associated with a 24 bp deletion in the *MC1R* gene, encoding a receptor with eight deleted amino acids (SNALETTI) at positions 87-94 (chapter 3). The wildtype grey phenotype was shown to be homozygous for the wildtype *MC1R-wt* allele, the jet black is homozygous for the *MC1RΔ24* allele and the brown-black was heterozygous (chapter 3). The *MC1RΔ24* allele appeared therefore to be incompletely dominant to the wildtype allele. Genetic tests revealed that the melanic grey squirrel found in the British Isles most probably originated from one or more introductions of melanic grey squirrels from North America (chapter 3). Melanism in the red and fox squirrel was found to be not associated with variations in the *MC1R* gene (chapter 3). Melanism was found to be associated with variations in the *ASIP* gene in the fox squirrel but not the red squirrel (chapter 4). Computer modelling predicted that the eight amino acid deletion in the MC1RΔ24 led to a receptor with a shorter ECL1 than the wildtype (chapter 5). *In vitro* experiments revealed that the eight amino acid deletion in the MC1RΔ24 led to a constitutively active receptor producing a higher basal level of intracellular cAMP compared to the wildtype receptor (chapter 6). Furthermore, α-MSH was found to act as an agonist to the MC1R-wt and MC1RΔ24 receptors whereas ASIP was found to act as an inverse agonist to the MC1R-wt and an agonist to the MC1RΔ24 receptor (chapter 6).

7.2 Genetic basis of melanism in the grey squirrel

The main aim of the research presented here was to understand the genetic and molecular basis of melanism in the grey squirrel. Three phenotypes of the grey squirrel were identified:

the wildtype grey, the brown-black and the jet black. Here, both the brown-black and jet black phenotypes are considered to be melanic. The first hypothesis was that variations in the *MC1R* gene are associated with melanism in the grey squirrel. This gene was the first candidate to investigate as the genetic basis of melanism because the MC1R is a key regulator of pigmentation, and variations in this gene have been associated with melanism in many species including mice (Robbins et al., 1993), foxes (Våge et al., 1997a) and sheep (Våge et al., 1999). In the present study, melanism was found to be associated with a 24 bp deletion in the *MC1R* of the grey squirrel, where the wildtype phenotype is homozygous for the *MC1R-wt* allele, the jet black phenotype is homozygous for the *MC1RΔ24* allele and the brown-black phenotype is heterozygous. The *MC1RΔ24* allele is thus incompletely dominant to the wildtype allele. The results of this part of the study have been published in the Journal of Heredity (McRobie, Thomas and Kelly, 2009) (see appendix 6).

The MC1R is a seven transmembrane GPCR which is predominantly expressed in melanocytes. This receptor has a high level of basal activity, which can be further activated by the agonist, α-MSH (Donatien et al., 1992; Holder and Haskell-Luevano, 2004). Activation of the MC1R leads to raised intracellular levels of cAMP by coupling to adenylate cyclase activity (Cone et al., 1993). High cAMP levels lead to activation of several signalling cascades including changes in gene expression via the cAMP responsive element binding protein (CREB) family of transcription factors. Raised cAMP levels have the overall effect of increased eumelanogenesis. A number of MC1R variants have been functionally characterised including mutants associated with melanism. It has been found that a number of these mutations cause the receptor to be constitutively active, leading to a higher level of

cAMP in the absence of stimulation, for example the E92K mutation in mice (Robbins et al., 1993), the C125R mutation in foxes (Våge et al., 1997a; Våge et al., 1999) and the M73K mutation in sheep (Våge et al., 1999). Given that these MC1R mutants are associated with melanism and that the mutations caused the receptors to be constitutively active, it was hypothesised that the deletion in the MC1R of the melanic grey squirrel leads to a constitutively active receptor. The results in the present study showed that the MC1RΔ24 did indeed have a higher basal level of activity compared to the MC1R-wt and so could be considered to be constitutively active. The results, as summarised in figure 6.7, show that both the wildtype and mutant receptors were found to be responsive to α-MSH. Stimulation with the inverse agonist ASIP reduced activity of the wildtype but increased activity of the mutant receptor. Overall, the MC1RΔ24 mutant produced consistently higher levels of intracellular cAMP, not only with basal levels of activity, but also resulting from both stimulation with α-MSH and stimulation with ASIP compared to the wildtype receptor. The results from this part of the study have been published in FEBS Letters (McRobie et al., 2014b) (see appendix 7).

7.3 The molecular basis of melanism in the grey squirrel

The eight amino acid deletion in the mutant MC1R was predicted to cause large overall architectural changes in the receptor. It was hypothesised that the close association between TM2 and TM3 together with ECL1 shortening could lead to an upwards shift in the TM2-TM3 bundle towards the extracellular side of the receptor. Such upward movement in GPCRs was thought to mimic that observed in the active receptor (Venkatakrishnan et al., 2013), and

this would be consistent with the finding that the MC1RΔ24 has a higher level of basal activity than the wildtype receptor.

The results presented here indicated that the E91 residue of the MC1R is of particular importance to the receptor's function. This residue is either deleted or substituted in a number of melanic variants of different species as well as the MC1RΔ24, as shown in table 3.8 (Robbins et al., 1993; Theron et al., 2001; Eizirik et al., 2003; Ling et al., 2003). The E91 contributes to an acidic pocket created by the E91, D114 and D118 residues (wildtype grey squirrel numbering) and the glutamic acid is thought to contribute to an electrostatic repulsion between TM2 and TM3. Removal or substitution of this residue is thought to mimic ligand binding, leading to activation of the receptor (Lu, Vage and Cone, 1998). Results presented in chapter 5 suggest that the MC1RΔ24 retains this acidic pocket even though E91 is deleted, with the computer models predicting that the E99 of the MC1R-wt is relocated to position 91 in the MC1RΔ24. In MCRs, this acidic pocket is thought to be essential for agonist binding (Lu, Vage and Cone, 1998) and this is consistent with the results found in the present study of both receptors being responsive to α-MSH. Results of mutagenesis studies on the MC1R-wt, presented in chapter 6 showed that the E91K mutant created a constitutively active receptor, further emphasising the importance of this residue.

7.4 ASIP acts as an agonist to the MC1RΔ24

The most exciting finding of this study was that ASIP acts as an agonist to the MC1RΔ24 although it is an inverse agonist to the MC1R-wt (McRobie et al., 2014b). This is the first

report of ASIP acting as an agonist to the MC1R for any species. In order to explain this result, it was noted that the eight amino acids deleted in the mutant receptor were predicted to be located in the second transmembrane domain of the receptor. This region of the receptor is a hot spot for other mutations associated with melanism, as shown in figure 1.16. Two of these mutations are deletions; the melanic jaguar has a five amino acid deletion and the melanic jaguarundi has an eight amino acid deletion, as shown in figure 1.17 (Eizirik et al., 2003). Two other species which have somewhat darkened phenotypes are the golden-headed lion tamarin, which has an eight amino acid deletion, and Eleonora's falcon, which has four amino acids deleted (Mundy and Kelly, 2006; Gangoso et al., 2011). Computer models predict that all of these deletions cause the ECL1 to be shortened in the mutant, as shown in figure 6.27. It has been shown that interactions between ECL1 of the MC1R and the C-terminus loop of ASIP are vital for ASIP to act as an inverse agonist and that ASIP with no C-terminus loop acts as an agonist (Patel et al., 2010). These authors predict that with interactions between the ECL1 and the C-terminus loop removed, ASIP is able to relocate to the same position as α-MSH and is able to activate the receptor as illustrated in figure 6.26 (Patel et al., 2010). A different study shows that the fragment peptide, CRFFNAFC, closely related to ASIP, (which includes the RFF motif of ASIP) also acts as an agonist to the MC1R (Haskell–Luevano et al., 2000). These authors also suggested that the peptide was acting in a similar way to α-MSH, with the RFF motif of ASIP mimicking the HFRW motif of α-MSH. A similar study showed that when the HFRW sequence of α-MSH is incorporated into a fragment of the agouti related protein (AgRP), which normally acts as an antagonist to the MC4R, it acts as an agonist (Jackson et al., 2005). These studies all show that the RFF sequence of ASIP and the HFRW sequence of α-MSH both interact with and activate the MC1R. It seems likely, therefore, that a similar event is occurring in the MC1R-wt and

MC1RΔ24. In the MC1R-wt, the ECL1 and C-terminus loop of ASIP interact and ASIP acts as an inverse agonist. In contrast, in the MC1RΔ24, the shortened ECL1 may lead to relocation of ASIP and may allow it to act as an agonist. This finding led to the prediction that a shorter ECL1 allows ASIP to function as an agonist. This prediction could be tested in further functional and mutational studies on the MC1R-wt (specifically altering ECL1), and also on functional studies on the MC1R of the jaguar, jaguarundi, golden-headed lion tamarin and Eleonora's falcon.

7.5 Genotype to phenotype

The finding that ASIP acted as an inverse agonist to the MC1R-wt but an agonist to the MC1RΔ24 was consistent with the phenotypic study of the grey squirrel. The wildtype was found to have hairs with wide bands of phaeomelanin (figure 3.3). These correspond to a pulse of ASIP expression during the hair growth cycle (Vrieling et al., 1994). The jet black squirrel, however, has no banding and the heterozygotes have a narrow band of phaeomelanin as discussed in chapter 3, suggesting that the switch from eu- to phaeomelanogenesis is delayed because of higher levels of cAMP compared to the wildtype. Furthermore, jet black squirrels do not have the lighter ventrum found in the other two morphs of the grey squirrel. In mice, as discussed in the introduction, ASIP is expressed continuously in the ventrum accounting for the lack of pigment in this region of the body (Vrieling et al., 1994). It has also been found that continuous stimulation with ASIP leads to dedifferentiation of melanoctyes (Le Pape et al., 2009). If this were also the case in the grey squirrel, this would account for the pale ventrum in the wildtype as little or no pigment is found in ventral hairs. In the case of the jet black squirrel, homozygous for the MC1RΔ24, stimulation by ASIP is

likely to have the opposite effect, since raising cAMP levels would lead to increased eumelanogenesis, leaving melanocytes fully differentiated. Finally, the orange ventrum of the heterozygotes is consistent with intermediate levels of cAMP leading to phaeomelanogenesis. These observations led to the prediction that continuous exposure to ASIP would lead to dedifferentiation in wildtype grey squirrel melanocytes homozygous for the MC1R-wt but to eumelanogenesis from differentiated melanocytes homozygous for the MC1RΔ24 and phaeomelanogenesis from differentiated heterozygous melanocytes. This prediction could be tested in further studies.

The results of the current study presented strong evidence that the eight amino acid deletion in the MC1RΔ24 is the cause of the melanic phenotype of the grey squirrel, as cells transfected with this receptor had raised intracellular levels of cAMP compared to cells transfected with the MC1R-wt. There are various ways that raised intracellular cAMP levels may lead to increased eumelanogenesis in the melanic grey squirrel. The most obvious and direct effect of high levels of cAMP in the cell is that genes required for eumelanogenesis (*TYR, TRP-1* and *DCT*) are up-regulated, leading to the production and deposition of eumelanin, as outlined in the introduction and summarised in figure 1.14. High levels of cAMP may have other effects: raised levels of cAMP (typically stimulated by the agonist α-MSH) lead to up-regulation of the *MC1R* gene in a positive feedback loop (Rouzaud et al., 2003). In mice it has been shown that there are in fact three different transcripts of the *MC1R.* Under normal conditions transcript 1 (T1) is produced by melanocytes which leads to constitutive expression of the MC1R (Rouzaud et al., 2003). However, higher levels of cAMP provoke a second transcript (T2) to be produced which is translated more efficiently

than T1 and results in a greater number of receptors being expressed on the surface of the cell. It was also found by Rouzaud et al. (2003) that when cells were stimulated with ASIP, which typically reduces cAMP levels, a third transcript (T3) was produced. Translation efficiency of T3 is found to be severely impaired resulting in a decrease in MC1R synthesis and expression (Rouzaud et al., 2003). In humans it has been found that raised cAMP levels lead to melanocyte proliferation and also increase melanocyte dendricity which is linked to greater deposition of melanins (Abdel-Malek et al., 1995). Raised cAMP levels in the melanocytes of melanic grey squirrels may also have these effects with increased cell proliferation, increased MC1R expression and increased dendricity, all of which would explain the melanic phenotype. Examination of epidermal tissues could determine whether there are any differences in morphology or number of melanocytes between the wildtype and melanic morphs of the grey squirrel.

7.6 The genetic basis of melanism in the fox and red squirrels

A secondary aim of this study was to understand the genetic basis of melanism in the fox and red squirrels. The results showed that melanism is not associated with mutations in the *MC1R* gene for either of these species. However, two alleles were identified in both species as detailed in chapter 3 and shown in table 3.6. The results of this part of the study have been published in the Journal of Heredity (McRobie et al., 2014b) (see appendix 8). In order to further investigate the genetic basis of melanism in the fox and red squirrel, a second candidate gene, *ASIP*, was examined. Variations in *ASIP* have been associated with melanism in a number of species including mice, foxes and leopards (Våge et al., 1997a; Kingsley et al., 2009; Schneider et al., 2012). All cases of melanism associated with *ASIP* are recessive,

loss of function mutations (Nachman, Hoekstra and D'Agostino, 2003). The study here showed that two amino acid substitutions are associated with melanism in the fox squirrel but no variations were associated with melanism in the red squirrel. Further studies on a larger sample size of fox squirrels will be needed to confirm the full details of these findings as melanic phenotypes in the fox squirrel are more varied than those of the grey squirrel. There may be complex epistatic interactions with the two identified alleles of the *MC1R* and the two mutations found in *ASIP* which would require a large scale investigation to unravel.

7.7 Convergent evolution

The genetic bases of melanism in the grey, fox and red squirrels appear to be different in each case, giving an excellent example of convergent evolution where different underlying genetic causes lead to the same phenotype. The study of melanism across many different species has provided multiple examples of convergent evolution; in some cases the *same* genes are associated with melanism in *distantly* related taxa and conversely *different* genes produce similar phenotypes in *closely* related taxa. Thus, similar phenotypes which fulfil the same ecological functions can be the result of different underlying molecular mechanisms (Gompel and Prud'homme, 2009; Manceau et al., 2010). For example, studies of white lizards have demonstrated that independent mutations on the *MC1R* are associated with similar phenotypes in three different species. The investigation by Rosenblum et al. (2010) revealed that one mutation affected MC1R signalling and a different mutation affected receptor integration into the membrane. Studies of deer mice by Kingsley et al., (2009) showed that both a 125kb deletion of exon 2 and a premature stop codon in exon 3 of *ASIP* produce similar dark phenotypes. Studies of beach mice by Hoekstra et al. (2006) showed that in one

population, a lighter phenotype is achieved by moving the dorsal-ventral boundary such that the lighter ventral area is enlarged, but in a different population of the same species, a lighter phenotype is the result of a R65C mutation in the MC1R. As well as coding regions of genes, mutations in *cis*-regulatory regions have been identified which contribute to lighter phenotypes; in the deer mouse a lighter phenotype is the result of increased levels of *ASIP* mRNA leading to a longer band of phaeomelanin in the hairs (Linnen et al., 2009). There are cases of the same gene and the same mutation in diverse species, for example the E92K mutation in MC1R in mice and birds as previously mentioned (Robbins et al., 1993; Theron et al., 2001). There are cases of the same gene, but functionally different mutations, for example MC1R and skin lightening in lizards. There are cases of the same gene with different but functionally equivalent mutations for example the E92K MC1R mutation in mice and the C125R MC1R mutation in foxes (Våge et al., 1997a). Finally, mutations in different genes may have the same ultimate outcome, for example the E92K in MC1R of mice and R96C in the ASIP of dogs (Kerns et al., 2004). In the case of the red squirrel, the melanic phenotypes are varied, giving a spectrum of colours from bright red through various shades of brown to black. Given the central importance of both the *MC1R* and *ASIP* genes in pigmentation variation, future studies into the genetic basis of melanism in the red squirrel should begin with an investigation into the regulatory regions of these genes. A spectrum of colours could be due to a number of regulatory differences in both genes interacting epistatically. Alternatively, given that there are over 100 loci involved in pigmentation, it seems possible that the genetic basis of melanism in the red squirrel could be due to a complex interaction of a few or indeed many genes.

7.8 Melanism and natural selection

Melanism occurs repeatedly across many diverse species from reptiles to birds to mammals (Rosenblum, Hoekstra and Nachman, 2004; Mundy, 2005). Melanin production requires energy and so is costly to an organism, therefore it follows that melanism may have an evolutionary advantage. Mutations leading to melanism associated with the MC1R are generally dominant (Robbins et al., 1993; Eizirik et al., 2003; Nachman, Hoekstra and D'Agostino, 2003) or incompletely dominant as is the case for the grey squirrel and jaguarundi (Eizirik et al., 2003; McRobie, Thomas and Kelly, 2009). Natural selection is able to work on these mutations in both homozygotes and heterozygotes as the phenotypes are visible in both. This is not the case with recessive mutations associated with ASIP. In rare, recessive mutations, natural selection is inefficient as the gene responsible is most often found in heterozygotes where the phenotype is hidden from selection. As Haldane argued in 1927 "it seems therefore doubtful whether natural selection in random mating organisms can cause the spread of autosomal recessive characters unless they are extraordinarily valuable to their possessors" (Haldane, 1927; Orr and Betancourt, 2001; Nachman, Hoekstra and D'Agostino, 2003). This is known as Haldane's sieve and it would suggest that melanism is "extraordinarily valuable" to the fox squirrel, if further studies reveal that it is indeed recessive. (It should be noted, however, that Haldane's sieve may be overcome with inbreeding.) It therefore follows that there is a selective advantage to melanism and one advantage could be crypsis, where melanic individuals are more camouflaged than their wildtype counterparts. This has been demonstrated for deer mice and rock pocket mice where dark individuals are found living on dark coloured substrates and vice versa (Nachman, Hoekstra and D'Agostino, 2003; Steiner, Weber and Hoekstra, 2007). It is not so clear what the advantage would be to a squirrel as the grizzled colouration of the wildtype grey squirrel

is well camouflaged against tree bark, as shown in figure 1.4. Indeed, the evidence presented in chapter 3 suggests that the *MC1R* of the wildtype is under purifying selection, which is consistent with the grizzled colouration being selectively advantageous. Furthermore, studies have confirmed that the wildtype morph of fox squirrels is more camouflaged against tree barks than the melanic morph (Kiltie, 1992a). Not only is the wildtype squirrel's dorsum well hidden against trees but the pale ventrum offers countershading which is thought to make an object appear flatter, thus providing further concealment (Thayer, 1896). One plausible advantage to the melanic morph is that it is better concealed against fire-burned trees. Indeed, studies have revealed that there is a positive correlation between wild fires and the frequency of melanism in the fox squirrel (Kiltie, 1992a; Kiltie, 1992b). Further studies on predation in fox squirrels revealed that red-tailed hawks responded more slowly to melanic models of fox squirrels when the models were moving, but more slowly to wildtype models if they were static (Kiltie, 1992b). Furthermore, melanism has been associated with burned vegetation in other species including moths and grasshoppers (Majerus, 1998) p33. Therefore, it follows that melanic morphs maybe more camouflaged than wildtype morphs in some circumstances and so have a selective advantage.

Camouflage does not provide perfect concealment, though, and grey squirrels fall prey to hawks in their native North America (Gurnell and Gurnell, 1987) p150. A plausible advantage to the melanic morph is that predators become accustomed to searching for a particular form. It has been shown that predators can find prey more efficiently by focussing on a "search image" (Croze, 1970). A hawk's visual attention may become honed to look for grizzled squirrels on tree bark. With their attention focussed on this "search image", the hawk

may miss a dark squirrel even though it may be objectively more conspicuous. Being hunted by search image can drive the evolution of polymorphisms as has been demonstrated in a number of studies including beach clams and guppies (Poulton, 1890; Croze, 1970; Endler, 1978; Whiteley, Owen and Smith, 1997; Bond, 2013). The benefits to the melanic squirrel of not matching a predator's search image may be slight but evolution works by marginal gains and the slight advantage may be seen after many generations. It has also been suggested that dark objects are simply more difficult to recognise than light objects, which may play a part in survival (Majerus, 1998).

7.9 The MC1R and pleiotropy

There may be other ways that melanism provides a selective advantage to the grey squirrel. Melanism may be subject to sex selection. A glossy black coat may be an indicator of good health (Jawor and Breitwisch, 2003) and therefore may be positively selected by mates (Andersson, 1994). Other reasons may be more related to physiology; studies on thermogenic capacity in the grey squirrel showed that melanic grey squirrels had 18% lower heat loss, 20% lower basal activity and 11% higher nonshivering thermogenesis capacity than wildtype squirrels (Ducharme, Larochelle and Richard, 1989). These findings seem to show that there may be an advantage to melanism in the grey squirrel. The mechanism for these findings is far from clear but it could be related to pleiotropy of the MC1R. The MC1R has been detected in a number of tissues including the Leydig cells of the testes which are involved in testosterone production (Thörnwall et al., 1997). Although there is no evidence currently, it may be hypothesised that a constitutively active MC1R expressed in these cells of male grey squirrels could affect levels of testosterone in the body which could then affect muscle mass

and behaviour. However, studies comparing the behaviour of melanic wildtype grey squirrel show no differences between the morphs (Gustafson and VanDruff, 1990). Also it has been suggested that MC1R protein detection has been over-reported in a number of studies due to the difficulties of antibody specificity in various techniques (Roberts, Newton and Sturm, 2007). There are a few reports where darkened phenotypes show some intriguing associations, for example social dominance and mating behaviour in sheep (Loehr et al., 2008), sexual selection and dominance in lions (West and Packer, 2002) and immunity in owls (Roulin and Ducrest, 2011). However, there are only two strong cases of MC1R involvement in a pleiotropic effect; MC1R in the nervous system mediates analgesia in female mice (Mogil et al., 2003) and loss of function mutations lead to a greater susceptibility to melanoma in humans (Beaumont et al., 2005; Pérez Oliva et al., 2009). Taken together, the evidence suggests that the MC1R has low pleiotropic effects and as such is a good target for coat colour evolution.

7.10 Why the melanocortin-1 receptor?

Evolution is not infinitely flexible and is constrained by the genetic and developmental mechanisms available. Natural selection is serendipitous, acting on variants randomly thrown up by mutations. The *MC1R* is repeatedly shown to be associated with pigmentation differences but, as previously noted, this may be due to ascertainment bias. However, it is a particularly good evolutionary target where small changes can lead to large phenotypic effects with low negative pleiotropic effects. The *MC1R* gene appears to be highly mutable with over 70 alleles identified in humans (Beaumont et al., 2005). This is likely to be due to relaxed selection pressure and not because the *MC1R* has a higher mutation rate than other

genes, indeed, synonymous substitution rates for this gene are equivalent to other genes (Mundy and Kelly, 2003; Gompel and Prud'homme, 2009). As previously noted, the MC1R has a high level of constitutive activity compared to other GPCRs (Holst and Schwartz, 2003). Typically, the MC1R is active due to its high level of flexibility, and it can be either further activated by an agonist or silenced by an inverse agonist. This is in contrast to other GPCRs which only have agonists and are far more bimodal, being either inactive or active, the most obvious example being rhodopsin which is switched completely "on" or completely "off" in response to light (Palczewski et al., 2000). This suggests that the MC1R is a robust molecule which may be relatively tolerant of mutations and still able to function to a greater or lesser degree. This may contribute to its central role in pigmentation variation as mutations in this gene may only affect the degree of its activity, either acting to increase or decrease activity of the receptor. This is demonstrated by the number of colour variations associated with amino acid variations in the receptor (see figure 1.16). Considering the relatively small size of the MC1R gene, it is a common evolutionary target for pigmentation variation and this may be because loss-of-function alleles still generate viable phenotypes. This may be different for ASIP, where loss-of-function may have more serious consequences. Melanism associated with variations in ASIP is caused by loss-of-function mutations. It seems far more likely that random mutations to a gene would cause loss-of-function and therefore it might be expected that melanism would be more frequently be associated with ASIP. There are a few possible reasons why this may not be the case; firstly, as has been mentioned, ascertainment bias may lead to over-representation of the MC1R in the literature, secondly, mutations in ASIP would be recessive and subjected to Haldane's sieve, as mentioned previously, and thirdly, mutations in ASIP may have a higher pleiotropic effect than the MC1R. ASIP has been detected in numerous tissues in humans, rabbits and cattle including the heart, liver,

kidneys, brain, adipose tissue, spleen, lungs, muscle and skin (Wilson et al., 1995; Girardot et al., 2005; Fontanesi et al., 2010). Such widespread expression implies that ASIP has more roles than are currently understood that go beyond regulation of pigmentation.

7.11 Mutational hotspots

It has been shown that the MC1R is a target for the evolution of pigmentation but there also appears to be a particular region *within* the gene which is a mutation target. Although mutations do occur in different regions of the receptor, as shown in figure 1.16, there is a concentration of mutations at the TM2-TM3 boundary, particularly for gain of function mutations. It could be that there are few mutational options for achieving a gain of function in GPCRs. This TM2-TM3 boundary is certainly a mutational hotspot for deletions and all known deletions of 6 bp or more in the *MC1R* are located here. This mutational hotspot may be similar to that in the *Pitx1* gene in sticklebacks. There is a particularly high prevalence of deletions in this gene and it is thought that there may be inherent structural features at the locus leading to these deletions (Chan et al., 2010). The deletions in the MC1R at the TM2-TM3 boundary include those previously mentioned associated with darkened phenotypes: the grey squirrel (24 bp) (McRobie, Thomas and Kelly, 2009), the jaguar (15 bp) and jaguarundi (24 bp) (Eizirik et al., 2003), the golden-headed lion tamarin (24 bp) (Mundy and Kelly, 2003) , Eleonora's falcon (12 bp) (Gangoso et al., 2011) and rabbits (6 bp) (Fontanesi et al., 2006). There are a number of other deletions in this region in other species which are not associated with melanism, for example in the rabbit (30 bp) (Fontanesi et al., 2006), and in various mustelid lineages; the wolverine *Gulo gulo* (15bp), five species of martens *Martes americana, M. martes, M. melampus and M. zibellina* (45 bp) and the stone marten *M. fonia*

which has a 10 bp insertion with a 28 bp deletion (Hosoda et al., 2005). For the mustelids, Hosoda et al., (2005) observed that two of the three deletion events were associated with hexanucleotide repeats at both ends of the deleted region, supporting a hypothesis of polymerase slippage (Nishizawa and Nishizawa, 2002). A similar pattern is found in the jaguarundi but not in the other species (Hosoda et al., 2005). Deletions may occur in other regions of the receptor but are not likely to be well tolerated.

The PSI-BLAST alignment in figure 5.2 shows that the MC1R-wt sequence on the top line has a longer ECL1 than all of the other GPCRs. A comparison of the number of amino acids of the MC1R-wt and those of the other GPCRs in this alignment shows that 2% have four fewer amino acids, 9% have five fewer amino acids, 81% have six fewer amino acids and 6% have seven fewer amino acids in their ECL1 than the MC1R-wt. This suggests that the ECL1 of the MC1R is unusually long for this group of GPCRs and this could account for the tolerance of deletions in this region. It is as though the MC1R has a bit of "slack" in this region and can afford to have amino acids deleted and still remain functional. This observation also suggests that during the evolution of the MC1R, an insertion at this region may have taken place which could account for the high degree of flexibility of this receptor. Considering all the deletions in this alignment, 94% of them are located in either the intra- or extracellular loops with only 6% occurring in the helices. This suggests that deletions in general are tolerated in these loops but not in the helices and is consistent with the observation that the transmembrane domains of GPCRs are highly conserved both from a sequence, and also from a structural point of view (Nygaard et al., 2009).

7.12 Effects of TM2 and TM3 deletions

It appears from the examples above that deletions in the TM2-TM3 boundary of the MC1R are sometimes, but not always, associated with melanic phenotypes. The 30 bp deletion in rabbits leads to a recessive red colour, as previously noted (Fontanesi et al., 2006). The results of functional studies with MC1Rdel2 and MC1Rdel4 presented in chapter 6 showed that these deletions caused loss-, not gain-of-function. The phenotypes of the mustelids with deletions are not considered to be melanic but they are certainly dark, generally ranging from dark to light brown (Hosoda et al., 2005). Further studies on functional characterisation of these receptors with deletions would be necessary to come to firm conclusions on the effects of deletions in the TM2-TM3 boundary on receptor function. It may be hypothesised that the deletions do indeed cause the receptors to be more active, which would lead to melanism, but that selection pressure to be more camouflaged has led to other genetic changes to mitigate the effects. For example, changes in regulatory regions could lead to reduced expression of the receptor to compensate for the effects of a highly active receptor. Indeed, the E94K (E92K equivalent) mutation in ruffed lemurs has been shown to lead to a constitutively active receptor, but these lemurs have phaeomelanic hair, demonstrating that constitutive activity does not necessarily lead to melanism (Haitina et al., 2007). This E94K receptor is still responsive to ASIP though, and it seems that ASIP expression levels are the major regulator of pigmentation in this case, where levels of ASIP compensate for the high level of activity in the receptor (Haitina et al., 2007). This finely balanced interaction between MC1R activity and ASIP is demonstrated in transgenic mice (Jackson et al., 2007). Here, human *MC1R* was expressed in a transgenic mouse model. The human MC1R was found to be expressed at lower levels than those of mice, but this was compensated for by increased sensitivity to α-MSH (Jackson et al., 2007). The human MC1R in the transgenic mice was found to respond

to ASIP, and mice were able to produce hairs with bands of phaeomelanin accordingly. If, however, the expression levels of the human MC1R were increased, the size of the phaeomelanin band decreased (Jackson et al., 2007) indicating the fine balance between MC1R and ASIP levels leading to different phenotypes. This interaction between MC1R and ASIP levels was seen earlier in the case of the beach mice, where increased ASIP expression led to wider bands of phaeomelanin and a lighter phenotype (Steiner, Weber and Hoekstra, 2007). These examples demonstrate that biochemical properties of the MC1R receptor do not necessarily predict colour phenotype and show how the final phenotype of an organism is the result of a finely tuned balance between functional activity of MC1R and expression levels of both the receptor and ASIP. The MC1R is a fascinating molecular switch at the heart of the adaptive pigmentation in a broad range of species. The genetics of pigmentation are complex and there are many avenues to explore.

7.13 Concluding remarks

The primary aim of this study was to investigate the genetic and molecular basis of melanism in the grey squirrel. The findings of this study revealed that melanism in the grey squirrel was associated with a 24 bp deletion in the *MC1R*. A secondary aim of the study was to investigation the genetic basis of melanism in the fox and red squirrels. The findings here revealed that there were no mutations in the *MC1R* gene associated with colour variations in these two species. Computer programmes predicted that the 24 bp deletion found in the *MC1R* of the grey squirrel led to alterations in the three dimensional structure of the receptor, leading to a shortened ECL1 in the mutant. Functional studies revealed that cells transfected with the *MC1RΔ24* gene had a higher basal level of activity than cells transfected with the

MC1R-wt gene. Both MC1R-wt and MC1RΔ24 cells could be further stimulated by α-MSH. However, while ASIP acted as an inverse agonist to the MC1R-wt, it acted as an agonist to the MC1RΔ24. This is the first report of ASIP acting as an agonist to the MC1R in any species. These findings provide the likely genetic and molecular cause of melanism in the grey squirrel.

7.14 Limitations

There are a number of limitations to this study. Firstly, the computer models created here are based on homology modelling and the best template available had a 30% identity with the MC1R amino acid sequence. Although this is considered to be an acceptable level of identity (Beuming and Sherman, 2012), it is far from ideal and therefore the three dimensional structure of the MC1R can only be an approximation at present. The docking programme used was unable to allow for the flexibility of the receptor. This flexibility is likely to be an important aspect of ligand binding and so the docking data is only a crude estimate of the actual interactions. As more crystal structures become available and as computational methods improve, this MC1R model and its interactions will become a more accurate prediction of the actual structure and behaviour of this molecule.

Secondly, the ASIP used in these experiments was ASIP [93-132] purchased from Phoenix peptides. This is the form of the protein used in all other studies in the literature (Sanchez-Mas et al., 2005), but it is not the full length native ASIP and specifically is not ASIP of the grey squirrel. While it is the closest approximation that is possible for the purposes of this

study, it was not ideal. Furthermore, the MC1R was transiently expressed in heterologous cells and not in melanocytes. Again, this is standard practice for studies of MC1R function presented in the literature (Jiménez-Cervantes et al., 2001), but ideally melanocytes would be obtained from both homozygous wildtype and melanic grey squirrels. Furthermore, in native tissue there may be differences in G-protein to receptor coupling and there may be other factors that affect ligand binding. The function of the accessory protein attractin *in vivo* may be affected by the eight amino acid deletion and endogenous α-MSH should also be taken into account. Despite these limitations, these *in vitro* results are consistent with the phenotypes observed. Future work using grey squirrel ASIP and grey squirrel melanocytes would address some of these limitations but at present these changes are not feasible.

7.15 Further work

The work in this study opens up many questions for further research in diverse directions, from looking at the structure of the MC1R molecule to the effect a molecular change may have on the behaviour of the whole animal. The most ambitious project would be to obtain the crystal structure of the MC1R-wt and MC1RΔ24 both in the active and inactive states. At present this is an unrealistic goal but listed below are a range of more realistic directions for further study. The suggestions for further work are organised into three four major areas; genetics, cell based studies and tissue based studies.

7.15.1 Genetic studies

Continue the investigation of the genetic basis of melanism in the fox squirrel by sequencing the coding region of both the *MC1R* and *ASIP* gene of a large sample size. This sample should include all the variations of phenotypes. The hypotheses are that melanism is associated with two mutations in exon 4 of *ASIP* and that different alleles of the *MC1R* and *ASIP* interact epistatically, resulting in a range of phenotypes in the fox squirrel.

Investigate the genetic basis of melanism in the red squirrel by sequencing exon 1 of the *ASIP* gene and regulatory regions of both the *MC1R* and *ASIP* genes. The hypothesis is that the genetic basis of melanism is associated with variation in regulatory regions of *MC1R* or *ASIP*.

Investigate variation in the *MC1R-wt* and *MC1RΔ24* alleles across North America and Britain. The hypothesis is that there is more variation in the *MC1R-wt* allele compared to the *MC1RΔ24* allele.

7.15.2 Cell based studies

Unless otherwise stated, the cell based studies proposed here would use the same materials and methods as described in this study.

Investigate the role of the hydrophobic region of ECL1 in ASIP interactions with the MC1R-wt. In order to investigate this, various mutants of the MC1R-wt would be created by

mutagenesis. These mutants could have the hydrophobic residues replaced with hydrophilic or polar residues or the hydrophobic residues could be removed completely. These mutants would then be functionally characterised by stimulation with ASIP. The hypothesis is that the hydrophobic residues of ECL1 in MC1R-wt are essential for ASIP to function as an inverse agonist.

Investigate the function of MC1R variants with deletions in the ECL1 from the jaguar, jaguarundi, golden headed lion tamarin, Eleonora's falcon, rabbit, and the three mustelid species with shortened ECL1s. This experiment would take the same form as the functional experiment investigating the MC1R-wt and MC1RΔ24 of the grey squirrel. The hypothesis is that a shortened ECL1 is associated with increased activity of the MC1R. This can also be extended to the MC1R of the red squirrel. The hypothesis is that the MC1R of the red squirrel has a lower basal level of activity than that of the MC1R-wt of the grey squirrel. It may be further hypothesised that the MC1R of the red squirrel is not responsive to α-MSH or ASIP as the red squirrel has solid red hairs with no banding.

Investigate the differences in binding of both NDP-MSH to the MC1R-wt and MC1RΔ24 receptors using competition binding assays with the radio ligand ^{125}NDP-MSH and unlabelled NDP-MSH. (NDP-MSH is a superpotent analogue of α-MSH and radio ligands for this are typically used for binding experiments.) The hypothesis is that there is no significant difference between binding of NDP-MSH between the MC1R-wt and MC1RΔ24 variants.

Investigate the differences in binding of the G-protein with both MC1R-wt and MC1RΔ24 using a time-resolved fluorescence-based GTP binding assay. The hypothesis is that G-proteins bind with greater affinity to the MC1RΔ24 than the MC1R-wt.

7.15.3 Tissue based studies

Using tissues obtained from wildtype and melanic grey squirrels, histological examination of the epidermal tissues could reveal the relative numbers and differences in morphology of melanocytes. The hypothesis is that there are fewer melanocytes and they are less dentritic in the wildtype compared to the melanic grey squirrel tissues.

Investigate expression of the MC1R in a range of tissues using immunohistochemistry, specifically those where ASIP and AgRP are expressed. This work would relate to the functional work using AgRP described earlier and be part of an investigation into pleiotropy and the MC1R. This investigation is problematic as it would require specific antibodies that would detect squirrel MC1R and also differentiate between the MCRs.

References

Abagyan, R., 2012. *ICM software manual.* Molsoft LLC, San Diego.

Abdel-Malek, Z., Swope, V.B., Suzuki, I., Akcali, C., Harriger, M.D., Boyce, S.T., Urabe, K. and Hearing, V.J., 1995. Mitogenic and melanogenic stimulation of normal human melanocytes by melanotropic peptides. *Proceedings of the National Academy of Sciences of the United States of America,* 92 (5), pp.1789-1793.

Abdel-Malek, Z.A., Scott, M.C., Furumura, M., Lamoreux, M.L., Ollmann, M., Barsh, G.S. and Hearing, V.J., 2001. The melanocortin 1 receptor is the principal mediator of the effects of agouti signaling protein on mammalian melanocytes. *Journal of Cell Science,* 114 (5), pp.1019-1024.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research,* 25 (17), pp.3389-3402.

Andersson, M.B., 1994. *Sexual selection.* Princeton University Press.

Aoki, H. and Moro, O., 2002. Involvement of microphthalmia-associated transcription factor (MITF) in expression of human melanocortin-1 receptor (MC1R). *Life Sciences,* 71 (18), pp.2171-2179.

Audet, M. and Bouvier, M., 2012. Restructuring G-protein-coupled receptor activation. *Cell,* 151 (1), pp.14-23.

Baião, P.C., Schreiber, E. and Parker, P.G., 2007. The genetic basis of the plumage polymorphism in red-footed boobies (*Sula sula*): a melanocortin-1 receptor (MC1R) analysis. *Journal of Heredity,* 98 (4), pp.287-292.

Baiao, P.C. and Parker, P.G., 2012. Evolution of the melanocortin-1 receptor (MC1R) in Boobies and Gannets (*Aves, Suliformes*). *Journal of Heredity,* 103 (3), pp.322-329.

Baker, D. and Sali, A., 2001. Protein structure prediction and structural genomics. *Science,* 294 (5540), pp.93-96.

Bar, I., Kaddar, E., Velan, A. and David, L., 2013. Melanocortin receptor 1 and black pigmentation in the Japanese ornamental carp (*Cyprinus carpio var. Koi*). *Frontiers in Genetics,* 4.

Barkalow, F.S. and Shorten, M., 1973. *World of the gray squirrel*.

Barrowclough, G.F. and Sibley, F.C., 1980. Feather pigmentation and abrasion: test of a hypothesis. *The Auk,* pp.881-883.

Beaumont, K.A., Newton, R.A., Smit, D.J., Leonard, J.H., Stow, J.L. and Sturm, R.A., 2005. Altered cell surface expression of human MC1R variant receptor alleles associated with red hair and skin cancer risk. *Human Molecular Genetics,* 14 (15), pp.2145-2154.

Benned-Jensen, T., Mokrosinski, J. and Rosenkilde, M.M., 2011. The E92K melanocortin 1 receptor mutant induces cAMP production and arrestin recruitment but not ERK activity indicating biased constitutive signaling. *PLOS ONE,* 6 (9), pp.e24644.

Bennett, D., 1993. Genetics, development, and malignancy of melanocytes. *International Review of Cytology,* 146, pp.191-260.

Bennett, D.C. and Lamoreux, M.L., 2003. The color loci of mice–a genetic century. *Pigment Cell Research,* 16 (4), pp.333-344.

Berryere, T.G., Kerns, J.A., Barsh, G.S. and Schmutz, S.M., 2005. Association of an Agouti allele with fawn or sable coat color in domestic dogs. *Mammalian Genome,* 16 (4), pp.262-272.

Bertolotto, C., Abbe, P., Hemesath, T.J., Bille, K., Fisher, D.E., Ortonne, J.P. and Ballotti, R., 1998. Microphthalmia gene product as a signal transducer in cAMP-induced differentiation of melanocytes. *Journal of Cell Biology,* 142 (3), pp.827-835.

Bertolotto, C., Bille, K., Ortonne, J.P. and Ballotti, R., 1996. Regulation of tyrosinase gene expression by cAMP in B16 melanoma cells involves two CATGTG motifs surrounding the TATA box: implication of the microphthalmia gene product. *Journal of Cell Biology,* 134 (3), pp.747-755.

Beuming, T. and Sherman, W., 2012. Current assessment of docking into GPCR crystal structures and homology models: successes, challenges, and guidelines. *Journal of Chemical Information and Modeling,* 52 (12), pp.3263-3277.

Black, C.C., 1995. Holarctic evolution and dispersal of squirrels (Rodentia: Sciuridae). 1995. *Evolutionary Biology.* Springer. pp.305-322.

Boecklen, W., 2014. *Biston betularia*. [online] Available at <http://web.nmsu.edu/~wboeckle/biston.html> [Accessed 15 June 2014].

Bokony, V., Liker, A., Szekely, T. and Kis, J., 2003. Melanin-based plumage coloration and flight displays in plovers and allies. *Proc. R. Soc. B,* 270 (1532), pp.2491-2497.

Bolin, K.A., Anderson, D.J., Trulson, J.A., Thompson, D.A., Wilken, J., Kent, S.B., Gantz, I. and Millhauser, G.L., 1999. NMR structure of a minimized human agouti related protein prepared by total chemical synthesis. *FEBS Letters,* 451 (2), pp.125-131.

Bond, A.B., 2013. *Concealing Coloration in Animals.* Harvard University Press.

Branden, C. and Tooze, J., 1991. *Introduction to Protein Structure.* Garland New York.

Buades, J.M., Rodríguez, V., Terrasa, B., Perez-Mellado, V., Brown, R.P., Castro, J.A., Picornell, A. and Ramon, M., 2013. Variability of the mc1r Gene in Melanic and Non-Melanic Podarcis lilfordi and Podarcis pityusensis from the Balearic Archipelago. *PLOS ONE,* 8 (1), pp.e53088.

Bultman, S.J., Michaud, E.J. and Woychik, R.P., 1992. Molecular characterization of the mouse agouti locus. *Cell,* 71 (7), pp.1195-1204.

Burtt Jr, E.H., 1986. An analysis of physical, physiological, and optical aspects of avian coloration with emphasis on wood-warblers. *Ornithological Monographs,* pp.iii-126.

Busca, R. and Ballotti, R., 2000. Cyclic AMP a key messenger in the regulation of skin pigmentation. *Pigment Cell Research,* 13 (2), pp.60-69.

Busca, R., Bertolotto, C., Abbe, P., Englaro, W., Ishizaki, T., Narumiya, S., Boquet, P., Ortonne, J.P. and Ballotti, R., 1998. Inhibition of Rho is required for cAMP-induced melanoma cell differentiation. *Molecular Biology of the Cell,* 9 (6), pp.1367-1378.

Cabrera-Vera, T.M., Vanhauwe, J., Thomas, T.O., Medkova, M., Preininger, A., Mazzoni, M.R. and Hamm, H.E., 2003. Insights into G protein structure, function, and regulation. *Endocrine Reviews,* 24 (6), pp.765-781.

Candille, S.I., Kaelin, C.B., Cattanach, B.M., Yu, B., Thompson, D.A., Nix, M.A., Kerns, J.A., Schmutz, S.M., Millhauser, G.L. and Barsh, G.S., 2007. A -defensin mutation causes black coat color in domestic dogs. *Science,* 318 (5855), pp.1418-1423.

Chai, B., Neubig, R.R., Millhauser, G.L., Thompson, D.A., Jackson, P.J., Barsh, G.S., Dickinson, C.J., Li, J., Lai, Y. and Gantz, I., 2003. Inverse agonist activity of agouti and agouti-related protein. *Peptides,* 24 (4), pp.603-609.

Chai, B., Pogozheva, I.D., Lai, Y., Li, J., Neubig, R.R., Mosberg, H.I. and Gantz, I., 2005. Receptor-antagonist interactions in the complexes of agouti and agouti-related protein with human melanocortin 1 and 4 receptors. *Biochemistry,* 44 (9), pp.3418-3431.

Chakraborty, A.K., Funasaka, Y., Slominski, A., Ermak, G., Hwang, J., Pawelek, J.M. and Ichihashi, M., 1996. Production and release of proopiomelanocortin (POMC) derived peptides by human melanocytes and keratinocytes in culture: regulation by ultraviolet B. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research,* 1313 (2), pp.130-138.

Chan, Y.F., Marks, M.E., Jones, F.C., Villarreal, G.,Jr, Shapiro, M.D., Brady, S.D., Southwick, A.M., Absher, D.M., Grimwood, J., Schmutz, J., Myers, R.M., Petrov, D., Jonsson, B., Schluter, D., Bell, M.A. and Kingsley, D.M., 2010. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. *Science,* 327 (5963), pp.302-305.

Chase, H.B., 1954. Growth of the hair. *Physiol Rev,* 34 (1), pp.113-126.

Cheviron, Z.A., Hackett, S.J. and Brumfield, R.T., 2006. Sequence variation in the coding region of the melanocortin-1 receptor gene (MC1R) is not associated with plumage variation in the blue-crowned manakin (Lepidothrix coronata). *Proc. R. Soc. B*, 273 (1594), pp.1613-1618.

Chhajlani, V. and Wikberg, J.E., 1992. Molecular cloning and expression of the human melanocyte stimulating hormone receptor cDNA. *FEBS Letters*, 309 (3), pp.417-420.

Ciampolini, R., Cecchi, F., Spaterna, A., Bramante, A., Bardet, S.M. and Oulmouden, A., 2013. Characterization of different 5′-untranslated exons of the ASIP gene in black-and-tan Doberman Pinscher and brindle Boxer dogs. *Animal Genetics*, 44 (1), pp.114-117.

Cone, R.D., Mountjoy, K.G., Robbins, L.S., Nadeau, J.H., Johnson, K.R., Roselli-Rehfuss, L. and Mortrud, M.T., 1993. Cloning and functional characterization of a family of receptors for the melanotropic peptides. *Annals of the New York Academy of Sciences*, 680 (1), pp.342-363.

Cook, L. and Saccheri, I., 2013. The peppered moth and industrial melanism: evolution of a natural selection case study. *Heredity*, 110 (3), pp.207-212.

Crick, F.H., 1953. The packing of-helices: simple coiled-coils. *Acta Crystallographica*, 6 (8-9), pp.689-697.

Croze, H., 1970. *Searching image in carrion crows: hunting strategy in a predator and some anti-predator devices in camouflaged prey.* Parey.

Darwin, E., 1796. Zoonomia, vol. 2. *London, J. Johnson*.

del Marmol, V. and Beermann, F., 1996. Tyrosinase and related proteins in mammalian pigmentation. *FEBS Letters*, 381 (3), pp.165-168.

Deupi, X. and Standfuss, J., 2011. Structural insights into agonist-induced activation of G-protein-coupled receptors. *Current Opinion in Structural Biology*, 21 (4), pp.541-551.

Donatien, P., Hunt, G., Pieron, C., Lunec, J., Taieb, A. and Thody, A., 1992. The expression of functional MSH receptors on cultured human melanocytes. *Archives of Dermatological Research*, 284 (7), pp.424-426.

Doré, A.S., Robertson, N., Errey, J.C., Ng, I., Hollenstein, K., Tehan, B., Hurrell, E., Bennett, K., Congreve, M. and Magnani, F., 2011. Structure of the Adenosine A< sub> 2A</sub> Receptor in Complex with ZM241385 and the Xanthines XAC and Caffeine. *Structure*, 19 (9), pp.1283-1293.

Dores, R.M. and Baron, A.J., 2011. Evolution of POMC: origin, phylogeny, posttranslational processing, and the melanocortins. *Annals of the New York Academy of Sciences*, 1220 (1), pp.34-48.

Drake, M.T., Shenoy, S.K. and Lefkowitz, R.J., 2006. Trafficking of G protein-coupled receptors. *Circulation Research,* 99 (6), pp.570-582.

Drögemüller, C., Giese, A., Martins-Wess, F., Wiedemann, S., Andersson, L., Brenig, B., Fries, R. and Leeb, T., 2006. The mutation causing the black-and-tan pigmentation phenotype of Mangalitza pigs maps to the porcine ASIP locus but does not affect its coding sequence. *Mammalian Genome,* 17 (1), pp.58-66.

Ducharme, M.B., Larochelle, J. and Richard, D., 1989. Thermogenic capacity in gray and black morphs of the gray squirrel, *Sciurus carolinensis*. *Physiological Zoology*, pp.1273-1292.

Ducrest, A.L., Amacker, M., Lingner, J. and Nabholz, M., 2002. Detection of promoter activity by flow cytometric analysis of GFP reporter expression. *Nucleic Acids Research,* 30 (14), pp.e65.

Duhl, D.M., Stevens, M.E., Vrieling, H., Saxon, P.J., Miller, M.W., Epstein, C.J. and Barsh, G.S., 1994. Pleiotropic effects of the mouse lethal yellow (Ay) mutation explained by deletion of a maternally expressed gene and the simultaneous production of agouti fusion RNAs. *Development,* 120 (6), pp.1695-1708.

Dymond, J.S., Scheifele, L.Z., Richardson, S., Lee, P., Chandrasegaran, S., Bader, J.S. and Boeke, J.D., 2009. Teaching synthetic biology, bioinformatics and engineering to undergraduates: the interdisciplinary Build-a-Genome course. *Genetics,* 181 (1), pp.13-21.

Eizirik, E., Yuhki, N., Johnson, W.E., Menotti-Raymond, M., Hannah, S.S. and O'Brien, S.J., 2003. Molecular genetics and evolution of melanism in the cat family. *Current Biology,* 13 (5), pp.448-453.

Ellis, L.S. and Maxson, L.R., 1980. Albumin evolution within new world squirrels (Sciuridae). *American Midland Naturalist,* pp.57-62.

Endler, J.A., 1978. A predator's view of animal color patterns. 1978. *Evolutionary Biology.* Springer, pp.319-364.

Everts, R., Rothuizen, J. and Oost, B., 2000. Identification of a premature stop codon in the melanocyte-stimulating hormone receptor gene (MC1R) in Labrador and Golden retrievers with yellow coat colour. *Animal Genetics,* 31 (3), pp.194-199.

Farrens, D.L., Altenbach, C., Yang, K., Hubbell, W.L. and Khorana, H.G., 1996. Requirement of rigid-body motion of transmembrane helices for light activation of rhodopsin. *Science,* 274 (5288), pp.768-770.

Fontanesi, L., Oulmouden, A., Tazzoli, M., Allain, D., Deretz-Picoulet, S., Robinson, T., Pecchioli, E., Cook, J., Russo, V. and Xicato, G., eds. 2008. *Proceedings of the 9th*

World Rabbit Congress, Verona, Italy, 10-13 June 2008. World Rabbit Science Association.

Fontanesi, L., Tazzoli, M., Beretti, F. and Russo, V., 2006. Mutations in the melanocortin 1 receptor (MC1R) gene are associated with coat colours in the domestic rabbit (Oryctolagus cuniculus). *Animal Genetics,* 37 (5), pp.489-493.

Fontanesi, L., Forestier, L., Allain, D., Scotti, E., Beretti, F., Deretz-Picoulet, S., Pecchioli, E., Vernesi, C., Robinson, T.J. and Malaney, J.L., 2010. Characterization of the rabbit agouti signaling protein (ASIP) gene: Transcripts and phylogenetic analyses and identification of the causative mutation of the nonagouti black coat colour. *Genomics,* 95 (3), pp.166-175.

Frändberg, P., Doufexis, M., Kapas, S. and Chhajlani, V., 2001. Cysteine residues are involved in structure and function of melanocortin 1 receptor: Substitution of a cysteine residue in transmembrane segment two converts an agonist to antagonist. *Biochemical and Biophysical Research Communications,* 281 (4), pp.851-857.

Fredriksson, R., Lagerstrom, M.C., Lundin, L.G. and Schioth, H.B., 2003. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Molecular Pharmacology,* 63 (6), pp.1256-1272.

Galandrin, S., Oligny-Longpré, G. and Bouvier, M., 2007. The evasive nature of drug efficacy: implications for drug discovery. *Trends in Pharmacological Sciences,* 28 (8), pp.423-430.

Gangoso, L., Grande, J.M., Ducrest, A., Figuerola, J., Bortolotti, G.R., Andrés, J. and Roulin, A., 2011. MC1R-dependent, melanin-based colour polymorphism is associated with cell-mediated response in the Eleonora's falcon. *Journal of Evolutionary Biology,* 24 (9), pp.2055-2063.

García-Borrón, J.C., Sánchez-Laorden, B.L. and Jiménez-Cervantes, C., 2005. Melanocortin-1 receptor structure and functional regulation. *Pigment Cell Research,* 18 (6), pp.393-410.

Gether, U., 2000. Uncovering molecular mechanisms involved in activation of G protein-coupled receptors. *Endocrine reviews,* 21 (1), pp.90-113.

Gether, U., Lin, S., Ghanouni, P., Ballesteros, J.A., Weinstein, H. and Kobilka, B.K., 1997. Agonists induce conformational changes in transmembrane domains III and VI of the β2 adrenoceptor. *The EMBO Journal,* 16 (22), pp.6737-6747.

Ghanouni, P., Gryczynski, Z., Steenhuis, J.J., Lee, T.W., Farrens, D.L., Lakowicz, J.R. and Kobilka, B.K., 2001. Functionally different agonists induce distinct conformations in the G protein coupling domain of the beta 2 adrenergic receptor. *The Journal of Biological Chemistry,* 276 (27), pp.24433-24436.

Girardot, M., Martin, J., Guibert, S., Leveziel, H., Julien, R. and Oulmouden, A., 2005. Widespread expression of the bovine Agouti gene results from at least three alternative promoters. *Pigment Cell Research,* 18 (1), pp.34-41.

Goding, C.R., 2000. Mitf from neural crest to melanoma: signal transduction and transcription in the melanocyte lineage. *Genes and Development,* 14 (14), pp.1712-1728.

Goldstein, G., Flory, K.R., Browne, B.A., Majid, S., Ichida, J.M., Burtt Jr, E.H. and Grubb Jr, T., 2004. Bacterial degradation of black and white feathers. *The Auk,* 121 (3), pp.656-659.

Gompel, N. and Prud'homme, B., 2009. The causes of repeated genetic evolution. *Developmental Biology,* 332 (1), pp.36-47.

Granier, S. and Kobilka, B., 2012. A new era of GPCR structural and chemical biology. *Nature Chemical Biology,* 8 (8), pp.670-673.

Granier, S., Manglik, A., Kruse, A.C., Kobilka, T.S., Thian, F.S., Weis, W.I. and Kobilka, B.K., 2012. Structure of the δ-opioid receptor bound to naltrindole. *Nature,* 485 (7398), pp.400-404.

Gratten, J., Pilkington, J., Brown, E., Beraldi, D., Pemberton, J. and Slate, J., 2010. The genetic basis of recessive self-colour pattern in a wild sheep population. *Heredity,* 104 (2), pp.206-214.

Gurnell, J. and Gurnell, J., 1987. *The natural history of squirrels.* Christopher Helm, Kent.

Gurnell, J., Wauters, L.A., Lurz, P.W. and Tosi, G., 2004. Alien species and interspecific competition: effects of introduced eastern grey squirrels on red squirrel population dynamics. *Journal of Animal Ecology,* 73 (1), pp.26-35.

Gustafson, E.J. and VanDruff, L.W., 1990. Behavior of black and gray morphs of *Sciurus carolinensis* in an urban environment. *American Midland Naturalist,* pp.186-192.

Haitina, T., Ringholm, A., Kelly, J., Mundy, N.I. and Schioth, H.B., 2007. High diversity in functional properties of melanocortin 1 receptor (MC1R) in divergent primate species is more strongly associated with phylogeny than coat color. *Molecular Biology and Evolution,* 24 (9), pp.2001-2008.

Haldane, J.B.S., ed. 1927. *Mathematical Proceedings of the Cambridge Philosophical Society.* Cambridge University Press.

Hamm, H.E., 1998. The many faces of G protein signaling. *The Journal of Biological Chemistry,* 273 (2), pp.669-672.

Haskell-Luevano, C., Cone, R.D., Monck, E.K. and Wan, Y., 2001. Structure activity studies of the melanocortin-4 receptor by in vitro mutagenesis: identification of agouti-related

protein (AGRP), melanocortin agonist and synthetic peptide antagonist interaction determinants. *Biochemistry,* 40 (20), pp.6164-6179.

Haskell–Luevano, C., Monck, E.K., Wan, Y. and Schentrup, A.M., 2000. The agouti-related protein decapeptide (Yc [CRFFNAFC] Y) possesses agonist activity at the murine melanocortin-1 receptor. *Peptides,* 21 (5), pp.683-689.

Haskell-Luevano, C., Sawyer, T.K., Hendrata, S., North, C., Panahinia, L., Stum, M., Staples, D.J., De Lauro Castrucci, Anna M, Hadley, M.E. and Hruby, V.J., 1996. Truncation studies of α-melanotropin peptides identify tripeptide analogues exhibiting prolonged agonist bioactivity. *Peptides,* 17 (6), pp.995-1002.

He, L., Gunn, T.M., Bouley, D.M., Lu, X., Watson, S.J., Schlossman, S.F., Duke-Cohan, J.S. and Barsh, G.S., 2001. A biochemical function for attractin in agouti-induced pigmentation and obesity. *Nature genetics,* 27 (1), pp.40-47.

Hearing, V.J., ed. 1999. *Journal of Investigative Dermatology Symposium Proceedings*. Nature Publishing Group.

Hearing, V.J. and Tsukamoto, K., 1991. Enzymatic control of pigmentation in mammals. *FASEB,* 5 (14), pp.2902-2909.

Hedegard, W. 2014. *Layers of the skin*. [online] Available at <http://cnx.org/content/m46060/latest/?collection=col11622/latest> [Accessed 13 July 2014].

Herraiz, C., Sánchez-Laorden, B.L., Jiménez-Cervantes, C. and García-Borrón, J.C., 2011. N-glycosylation of the human melanocortin 1 receptor: occupancy of glycosylation sequons and functional role. *Pigment Cell and Melanoma Research,* 24 (3), pp.479-489.

Herron, M.D., Castoe, T.A. and Parkinson, C.L., 2004. Sciurid phylogeny and the paraphyly of Holarctic ground squirrels (Spermophilus). *Molecular Phylogenetics and Evolution,* 31 (3), pp.1015-1030.

Hida, T., Wakamatsu, K., Sviderskaya, E.V., Donkin, A.J., Montoliu, L., Lynn Lamoreux, M., Yu, B., Millhauser, G.L., Ito, S. and Barsh, G.S., 2009. Agouti protein, mahogunin, and attractin in pheomelanogenesis and melanoblast-like alteration of melanocytes: a cAMP-independent pathway. *Pigment Cell and Melanoma Research*, 22 (5), pp.623-634.

Hodgkinson, C.A., Moore, K.J., Nakayama, A., Steingrímsson, E., Copeland, N.G., Jenkins, N.A. and Arnheiter, H., 1993. Mutations at the mouse microphthalmia locus are associated with defects in a gene encoding a novel basic-helix-loop-helix-zipper protein. *Cell,* 74 (2), pp.395-404.

Hoekstra, H.E., Hirschmann, R.J., Bundey, R.A., Insel, P.A. and Crossland, J.P., 2006. A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science,* 313 (5783), pp.101-104.

Holder, J.R. and Haskell-Luevano, C., 2004. Melanocortin ligands: 30 years of structure–activity relationship (SAR) studies. *Medicinal Research Reviews,* 24 (3), pp.325-356.

Holst, B. and Schwartz, T.W., 2003. Molecular mechanism of agonism and inverse agonism in the melanocortin receptors. *Annals of the New York Academy of Sciences,* 994 (1), pp.1-11.

Holst, B., Elling, C.E. and Schwartz, T.W., 2002. Metal ion-mediated agonism and agonist enhancement in melanocortin MC1 and MC4 receptors. *The Journal of Biological Chemistry,* 277 (49), pp.47662-47670.

Hosoda, T., Sato, J.J., Shimada, T., Campbell, K.L. and Suzuki, H., 2005. Independent nonframeshift deletions in the MC1R gene are not associated with melanistic coat coloration in three mustelid lineages. *Journal of Heredity,* 96 (5), pp.607-613.

Hughes, M.J., Lingrel, J.B., Krakowsky, J.M. and Anderson, K.P., 1993. A helix-loop-helix transcription factor-like gene is located at the mi locus. *The Journal of Biological Chemistry,* 268 (28), pp.20687-20690.

Jackson, P.J., Douglas, N.R., Chai, B., Binkley, J., Sidow, A., Barsh, G.S. and Millhauser, G.L., 2006. Structural and molecular evolutionary analysis of Agouti and Agouti-related proteins. *Chemistry and Biology,* 13 (12), pp.1297-1305.

Jackson, P.J., Yu, B., Hunrichs, B., Thompson, D.A., Chai, B., Gantz, I. and Millhauser, G.L., 2005. Chimeras of the agouti-related protein: Insights into agonist and antagonist selectivity of melanocortin receptors. *Peptides,* 26 (10), pp.1978-1987.

Jackson, I.J., Budd, P.S., Keighren, M. and McKie, L., 2007. Humanized MC1R transgenic mice reveal human specific receptor function. *Human Molecular Genetics,* 16 (19), pp.2341-2348.

Jaga, J., 2008. *Sciurus carolinensis*. [online] Available at <http://en.wikipedia.org/wiki/Black_squirrel> [Accessed 14 May 2014].

Jawor, J.M. and Breitwisch, R., 2003. Melanin ornaments, honesty, and sexual selection. *The Auk,* 120 (2), pp.249-265.

Jiménez-Cervantes, C., Germer, S., González, P., Sánchez, J., Sánchez, C.O. and García-Borrón, J.C., 2001. Thr40 and Met122 are new partial loss-of-function natural mutations of the human melanocortin 1 receptor. *FEBS Letters,* 508 (1), pp.44-48.

Johnson, J.A., Ambers, A.D. and Burnham, K.K., 2012. Genetics of plumage color in the gyrfalcon (Falco rusticolus): analysis of the melanocortin-1 receptor gene. *Journal of Heredity,* 103 (3), pp.315-321.

Kaczanowski, S. and Zielenkiewicz, P., 2010. Why similar protein sequences encode similar three-dimensional structures. *Theoretical Chemistry Accounts,* 125 (3-6), pp.643-650.

Katritch, V., Cherezov, V. and Stevens, R.C., 2012. Diversity and modularity of G protein-coupled receptor structures. *Trends in pharmacological sciences,* 33 (1), pp.17-27.

Kelley, L.A. and Sternberg, M.J., 2009. Protein structure prediction on the Web: a case study using the Phyre server. *Nature protocols,* 4 (3), pp.363-371.

Kerje, S., Lind, J., Schütz, K., Jensen, P. and Andersson, L., 2003. Melanocortin 1-receptor (MC1R) mutations are associated with plumage colour in chicken. *Animal Genetics,* 34 (4), pp.241-248.

Kerns, J.A., Newton, J., Berryere, T.G., Rubin, E.M., Cheng, J., Schmutz, S.M. and Barsh, G.S., 2004. Characterization of the dog Agouti gene and a nonagoutimutation in German Shepherd Dogs. *Mammalian Genome,* 15 (10), pp.798-808.

Kettlewell, H.B.D., 1955. Selection experiments on industrial melanism in the Lepidoptera. *Heredity,* 9 (3), pp.323-342.

Kijas, J.M., Wales, R., Tornsten, A., Chardon, P., Moller, M. and Andersson, L., 1998. Melanocortin receptor 1 (MC1R) mutations and coat color in pigs. *Genetics,* 150 (3), pp.1177-1185.

Kiltie, R.A., 1992a. Camouflage comparisons among fox squirrels from the Mississippi river delta. *Journal of Mammalogy,* pp.906-913.

Kiltie, R.A., 1992b. Tests of hypotheses on predation as a factor maintaining polymorphic melanism in coastal-plain fox squirrels (*Sciurus niger L.*). *Biological Journal of the Linnean Society,* 45 (1), pp.17-37.

Kingsley, E.P., Manceau, M., Wiley, C.D. and Hoekstra, H.E., 2009. Melanism in Peromyscus is caused by independent mutations in Agouti. *PLOS ONE,* 4 (7), pp.e6435.

Kjelsberg, M.A., Cotecchia, S., Ostrowski, J., Caron, M.G. and Lefkowitz, R.J., 1992. Constitutive activation of the alpha 1B-adrenergic receptor by all amino acid substitutions at a single site. Evidence for a region which constrains receptor activation. *The Journal of Biological Chemistry,* 267 (3), pp.1430-1433.

Klebig, M.L., Wilkinson, J.E., Geisler, J.G. and Woychik, R.P., 1995. Ectopic expression of the agouti gene in transgenic mice causes obesity, features of type II diabetes, and yellow fur. *Proceedings of the National Academy of Sciences of the United States of America,* 92 (11), pp.4728-4732.

Klungland, H., Vage, D., Gomez-Raya, L., Adalsteinsson, S. and Lien, S., 1995. The role of melanocyte-stimulating hormone (MSH) receptor in bovine coat color determination. *Mammalian Genome,* 6 (9), pp.636-639.

Knall, C. and Johnson, G.L., 1998. G-protein regulatory pathways: Rocketing into the twenty-first century. *Journal of Cellular Biochemistry,* 72 (S30–31), pp.137-146.

Kobilka, B.K., 2007. G protein coupled receptor structure and activation. *Biochimica et Biophysica Acta (BBA)-Biomembranes,* 1768 (4), pp.794-807.

Kruse, A.C., Ring, A.M., Manglik, A., Hu, J., Hu, K., Eitel, K., Hübner, H., Pardon, E., Valant, C. and Sexton, P.M., 2013. Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature,* 504 (7478), pp.101-106.

Kubic, J.D., Young, K.P., Plummer, R.S., Ludvik, A.E. and Lang, D., 2008. Pigmentation PAX-ways: the role of Pax3 in melanogenesis, melanocyte stem cell maintenance, and disease. *Pigment Cell and Melanoma Research,* 21 (6), pp.627-645.

Kufareva, I., Rueda, M., Katritch, V., Stevens, R.C. and Abagyan, R., 2011. Status of GPCR modeling and docking as reflected by community-wide GPCR Dock 2010 assessment. *Structure,* 19 (8), pp.1108-1126.

Kuramoto, T., Nomoto, T., Sugimura, T. and Ushijima, T., 2001. Cloning of the rat agouti gene and identification of the rat nonagouti mutation. *Mammalian Genome,* 12 (6), pp.469-471.

Kwon, H.Y., Bultman, S.J., Loffler, C., Chen, W.J., Furdon, P.J., Powell, J.G., Usala, A.L., Wilkison, W., Hansmann, I. and Woychik, R.P., 1994. Molecular structure and chromosomal mapping of the human homolog of the agouti gene. *Proceedings of the National Academy of Sciences of the United States of America,* 91 (21), pp.9760-9764.

Le Guilloux, V., Schmidtke, P. and Tuffery, P., 2009. Fpocket: an open source platform for ligand pocket detection. *BMC Bioinformatics,* 10, pp.168-2105-10-168.

Le Pape, E., Passeron, T., Giubellino, A., Valencia, J.C., Wolber, R. and Hearing, V.J., 2009. Microarray analysis sheds light on the dedifferentiating role of agouti signal protein in murine melanocytes via the Mc1r. *Proceedings of the National Academy of Sciences of the United States of America,* 106 (6), pp.1802-1807.

Lebon, G., Warne, T., Edwards, P.C., Bennett, K., Langmead, C.J., Leslie, A.G. and Tate, C.G., 2011. Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation. *Nature,* 474 (7352), pp.521-525.

Lee, C., Yun, J., Lim, S. and Lee, W., 2010. Solution structures and molecular interactions of selective melanocortin receptor antagonists. *Molecules and Cells,* 30 (6), pp.551-556.

Leeb, T., Deppe, A., Kriegesmann, B. and Brenig, B., 2000. Genomic structure and nucleotide polymorphisms of the porcine agouti signalling protein gene (ASIP). *Animal Genetics,* 31 (5), pp.335-335.

Lefkowitz, R.J. and Shenoy, S.K., 2005. Transduction of receptor signals by beta-arrestins. *Science,* 308 (5721), pp.512-517.

Lerner, A.B., 1993. The Discovery of the Melanotropins: A History of Pituitary Endocrinologya. *Annals of the New York Academy of Sciences,* 680 (1), pp.1-12.

Lerner, A.B. and McGuire, J.S., 1961. Effect of alpha- and betamelanocyte stimulating hormones on the skin colour of man. *Nature,* 189, pp.176-179.

Lin, J.Y. and Fisher, D.E., 2007. Melanocyte biology and skin pigmentation. *Nature,* 445 (7130), pp.843-850.

Ling, M.K., Lagerström, M.C., Fredriksson, R., Okimoto, R., Mundy, N.I., Takeuchi, S. and Schiöth, H.B., 2003. Association of feather colour with constitutively active melanocortin 1 receptors in chicken. *European Journal of Biochemistry,* 270 (7), pp.1441-1449.

Linnen, C.R., Kingsley, E.P., Jensen, J.D. and Hoekstra, H.E., 2009. On the origin and spread of an adaptive allele in deer mice. *Science,* 325 (5944), pp.1095-1098.

Loehr, J., Carey, J., Ylönen, H. and Suhonen, J., 2008. Coat darkness is associated with social dominance and mating behaviour in a mountain sheep hybrid lineage. *Animal Behaviour,* 76 (5), pp.1545-1553.

Loser, I., 2007. *Sciurus niger*. [online] Available at < http://commons.wikimedia.org/wiki/File:Sciurus_niger.JPG> [Accessed 12 May 2014].

Lu, D., Vage, D.I. and Cone, R.D., 1998. A ligand-mimetic model for constitutive activation of the melanocortin-1 receptor. *Molecular Endocrinology,* 12 (4), pp.592-604.

Lu, D., Willard, D., Patel, I.R., Kadwell, S., Overton, L., Kost, T., Luther, M., Chen, W., Woychik, R.P. and Wilkison, W.O., 1994. Agouti protein is an antagonist of the melanocyte-stimulating-hormone receptor. *Nature,* 371 (6500) pp.799-802.

Luttrell, L.M. and Lefkowitz, R.J., 2002. The role of beta-arrestins in the termination and transduction of G-protein-coupled receptor signals. *Journal of Cell Science,* 115 (Pt 3), pp.455-465.

MacDougall-Shackleton, E.A., Blanchard, L., Igdoura, S.A. and Gibbs, H.L., 2003. Unmelanized plumage patterns in Old World leaf warblers do not correspond to sequence variation at the melanocortin-1 receptor locus (MC1R). *Molecular Biology and Evolution,* 20 (10), pp.1675-1681.

Macindoe, G., Mavridis, L., Venkatraman, V., Devignes, M.D. and Ritchie, D.W., 2010. HexServer: an FFT-based protein docking server powered by graphics processors. *Nucleic Acids Research,* 38 (Web Server issue), pp.W445-9.

Madabushi, S., Gross, A.K., Philippi, A., Meng, E.C., Wensel, T.G. and Lichtarge, O., 2004. Evolutionary trace of G protein-coupled receptors reveals clusters of residues that

determine global and class-specific functions. *The Journal of Biological Chemistry,* 279 (9), pp.8126-8132.

Majerus, M., Brunton, C. and Stalker, J., 2000. A bird's eye view of the peppered moth. *Journal of Evolutionary Biology,* 13 (2), pp.155-159.

Majerus, M.E., 1998. *Melanism: Evolution in Action.* Oxford University Press Oxford.

Majerus, M.E. and Mundy, N.I., 2003. Mammalian melanism: natural selection in black and white. *Trends in Genetics,* 19 (11), pp.585-588.

Manceau, M., Domingues, V.S., Linnen, C.R., Rosenblum, E.B. and Hoekstra, H.E., 2010. Convergence in pigmentation at multiple levels: mutations, genes and function. *Philosophical transactions of the Royal Society of London.Series B, Biological sciences,* 365 (1552), pp.2439-2450.

Manceau, M., Domingues, V.S., Mallarino, R. and Hoekstra, H.E., 2011. The developmental role of Agouti in color pattern evolution. *Science,* 331 (6020), pp.1062-1065.

Marieb, E.N. and Hoehn, K., 2007. *Human anatomy and physiology.* Pearson Education.

Marklund, L., Moller, M.J., Sandberg, K. and Andersson, L., 1996. A missense mutation in the gene for melanocyte-stimulating hormone receptor (MCIR) is associated with the chestnut coat color in horses. *Mammalian Genome,* 7 (12), pp.895-899.

Más, J.S., Sánchez, C.O., Ghanem, G., Haycock, J., Teruel, J.A.L., García-Borrón, J.C. and Jiménez-Cervantes, C., 2002. Loss-of-function variants of the human melanocortin-1 receptor gene in melanoma cells define structural determinants of receptor function. *European Journal of Biochemistry,* 269 (24), pp.6133-6141.

Mayer, T.C., 1973. The migratory pathway of neural crest cells into the skin of mouse embryos. *Developmental biology,* 34 (1), pp.39-46.

McNulty, J.C., Jackson, P.J., Thompson, D.A., Chai, B., Gantz, I., Barsh, G.S., Dawson, P.E. and Millhauser, G.L., 2005. Structures of the agouti signaling protein. *Journal of Molecular Biology,* 346 (4), pp.1059-1070.

McRobie, H.R., 2014a. *The Black Squirrel Project*. [on-line] Available at: <www.blacksquirrelproject.org> [Accessed 25 July 2014].

McRobie, H.R., King, L.M., Fanutti, C., Symmons, M.F. and Coussons, P.J., 2014b. Agouti signalling protein is an inverse agonist to the wildtype and agonist to the melanic variant of the melanocortin-1 receptor in the grey squirrel (*Sciurus carolinensis*). *FEBS Letters, 588 (14), pp. 2335-2343.*

McRobie, H., Thomas, A. and Kelly, J., 2009. The genetic basis of melanism in the gray squirrel (*Sciurus carolinensis*). *Journal of Heredity,* 100 (6), pp.709-714.

McRobie, H.R., King, L.M., Fanutti, C., Coussons, P.J., Moncrief, N.D. and Thomas, A.P., 2014. Melanocortin 1 receptor (MC1R) gene sequence variation and melanism in the gray (*Sciurus carolinensis*), fox (*Sciurus niger*), and red (*Sciurus vulgaris*) squirrel. *Journal of Heredity,* 105 (3), pp.423-428.

Michino, M., Abola, E., GPCR Dock 2008 participants, Brooks, C.L.,3rd, Dixon, J.S., Moult, J. and Stevens, R.C., 2009. Community-wide assessment of GPCR structure modelling and ligand docking: GPCR Dock 2008. *Nature Reviews.Drug Discovery,* 8 (6), pp.455-463.

Middleton, A.D., 1931. *The grey squirrel: the introduction and spread of the American grey squirrel in the British Isles, its habits, food, and relations with the native fauna of the country.* Sidgwick and Jackson, Ltd.

Miller, M.W., Duhl, D.M., Vrieling, H., Cordes, S.P., Ollmann, M.M., Winkes, B.M. and Barsh, G.S., 1993. Cloning of the mouse agouti gene predicts a secreted protein ubiquitously expressed in mice carrying the lethal yellow mutation. *Genes and Development,* 7 (3), pp.454-467.

Miltenberger, R.J., Wakamatsu, K., Ito, S., Woychik, R.P., Russell, L.B. and Michaud, E.J., 2002. Molecular and phenotypic analysis of 25 recessive, homozygous-viable alleles at the mouse agouti locus. *Genetics,* 160 (2), pp.659-674.

Minvielle, F., Gourichon, D., Ito, S. Inoue-Murayama, M. and Rivière, S., 2007. Effects of the dominant lethal *yellow* mutation on reproduction, growth, feed consumption, body temperature, and body composition in Japanese quail. *Poult. Sci*. 86 pp. 1646–1650.

Mogil, J.S., Wilson, S.G., Chesler, E.J., Rankin, A.L., Nemmani, K.V., Lariviere, W.R., Groce, M.K., Wallace, M.R., Kaplan, L., Staud, R., Ness, T.J., Glover, T.L., Stankova, M., Mayorov, A., Hruby, V.J., Grisel, J.E. and Fillingim, R.B., 2003. The melanocortin-1 receptor gene mediates female-specific mechanisms of analgesia in mice and humans. *Proceedings of the National Academy of Sciences of the United States of America,* 100 (8), pp.4867-4872.

Moncrief, N.D., Lack, J.B. and Van Den Bussche, Ronald A, 2010. Eastern fox squirrel (*Sciurus niger*) lacks phylogeographic structure: recent range expansion and phenotypic differentiation. *Journal of Mammalogy*, 91 (5), pp.1112-1123.

Mountjoy, K.G., Robbins, L.S., Mortrud, M.T. and Cone, R.D., 1992. The cloning of a family of genes that encode the melanocortin receptors. *Science,* 257 (5074), pp.1248-1251.

Mundy, N. and Kelly, J., 2003. Evolution of a pigmentation gene, the melanocortin-1 receptor, in primates. *American Journal of Physical Anthropology,* 121 (1), pp.67-80.

Mundy, N.I., Badcock, N.S., Hart, T., Scribner, K., Janssen, K. and Nadeau, N.J., 2004. Conserved genetic basis of a quantitative plumage trait involved in mate choice. *Science*, 303 (5665), pp.1870-1873.

Mundy, N.I. and Kelly, J., 2006. Investigation of the role of the agouti signaling protein gene (ASIP) in coat color evolution in primates. *Mammalian Genome*, 17 (12), pp.1205-1213.

Mundy, N.I., 2005. A window on the genetics of evolution: MC1R and plumage colouration in birds. *Proc. R. Soc. B,* 272 (1573), pp.1633-1640.

Nachman, M.W., Hoekstra, H.E. and D'Agostino, S.L., 2003. The genetic basis of adaptive melanism in pocket mice. *Proceedings of the National Academy of Sciences of the United States of America*, 100 (9), pp.5268-5273.

Nadeau, N., Minvielle, F. and Mundy, N., 2006. Association of a Glu92Lys substitution in MC1R with extended brown in Japanese quail (*Coturnix japonica*). *Animal Genetics*, 37 (3), pp.287-289.

Nadeau, N., Minvielle, F., Ito, S., Inoue-Murayama, M., Gourichon, D., Follett, S., Burke, T. and Mundy, N., 2008. Characterization of Japanese quail *yellow* as a genomic deletion upstream of the avian homolog of the mammalian *ASIP* (*agouti*) gene. *Genetics*, 178, pp. 777-786.

Nishimura, E.K., Jordan, S.A., Oshima, H., Yoshida, H., Osawa, M., Moriyama, M., Jackson, I.J., Barrandon, Y., Miyachi, Y. and Nishikawa, S., 2002. Dominant role of the niche in melanocyte stem-cell fate determination. *Nature*, 416 (6883), pp.854-860.

Nishizawa, M. and Nishizawa, K., 2002. A DNA sequence evolution analysis generalized by simulation and the Markov chain Monte Carlo method implicates strand slippage in a majority of insertions and deletions. *Journal of Molecular Evolution*, 55 (6), pp.706-717.

Norris, B.J. and Whan, V.A., 2008. A gene duplication affecting expression of the ovine ASIP gene is responsible for white and black sheep. *Genome Research*, 18 (8), pp.1282-1293.

Nygaard, R., Frimurer, T.M., Holst, B., Rosenkilde, M.M. and Schwartz, T.W., 2009. Ligand binding and micro-switches in 7TM receptor structures. *Trends in Pharmacological Sciences*, 30 (5), pp.249-259.

Orr, H.A. and Betancourt, A.J., 2001. Haldane's sieve and adaptation from the standing genetic variation. *Genetics*, 157 (2), pp.875-884.

Orry, A.J. and Abagyan, R., 2012. *Homology modeling: ethods and protocols.* Humana Press.

Ortonne, J. and Prota, G., 1993. Hair melanins and hair color: ultrastructural and biochemical aspects. *Journal of Investigative Dermatology*, 101, pp.82S-89S.

Øyehaug, L., Plahte, E., Vage, D. and Omholt, S., 2002. The regulatory basis of melanogenic switching. *Journal of Theoretical Biology*, 215 (4), pp.449-468.

Ozeki, H., Ito, S., Wakamatsu, K. and Hirobe, T., 1995. Chemical characterization of hair melanins in various coat-color mutants of mice. *Journal of Investigative Dermatology*, 105 (3).

Palczewski, K., Kumasaka, T., Hori, T., Behnke, C.A., Motoshima, H., Fox, B.A., Le Trong, I., Teller, D.C., Okada, T., Stenkamp, R.E., Yamamoto, M. and Miyano, M., 2000. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science*, 289 (5480), pp.739-745.

Parnot, C., Miserey-Lenkei, S., Bardin, S., Corvol, P. and Clauser, E., 2002. Lessons from constitutively active mutants of G protein-coupled receptors. *Trends in Endocrinology and Metabolism*, 13 (8), pp.336-343.

Patel, M.P., Cribb Fabersunne, C.S., Yang, Y., Kaelin, C.B., Barsh, G.S. and Millhauser, G.L., 2010. Loop-swapped chimeras of the agouti-related protein and the agouti signaling protein identify contacts required for melanocortin 1 receptor selectivity and antagonism. *Journal of Molecular Biology*, 404 (1), pp.45-55.

Pawelek, J.M., 1979. Evidence suggesting that a cyclic AMP-dependent protein kinase is a positive regulator of proliferation in Cloudman S91 melanoma cells. *Journal of Cellular Physiology*, 98 (3), pp.619-625.

Pawelek, J., 1991. After dopachrome? *Pigment Cell Research*, 4 (2), pp.53-62.

Pawelek, J.M., 1985. Studies on the Cloudman melanoma cell line as a model for the action of MSH. *The Yale Journal of Biology and Medicine*, 58 (6), pp.571-578.

Pérez Oliva, A.B., Fernéndez, L.P., DeTorre, C., Herráiz, C., Martínez-Escribano, J.A., Benítez, J., Lozano Teruel, J.A., García-Borrón, J.C., Jiménez-Cervantes, C. and Ribas, G., 2009. Identification and functional analysis of novel variants of the human melanocortin 1 receptor found in melanoma patients. *Human Mutation*, 30 (5), pp.811-822.

Perry, W.L., Nakamura, T., Swing, D.A., Secrest, L., Eagleson, B., Hustad, C.M., Copeland, N.G. and Jenkins, N.A., 1996. Coupled site-directed mutagenesis/transgenesis identifies important functional domains of the mouse agouti protein. *Genetics*, 144 (1), pp.255-264.

Pitcher, J.A., Freedman, N.J. and Lefkowitz, R.J., 1998. G protein-coupled receptor kinases. *Annual Review of Biochemistry*, 67 (1), pp.653-692.

Pointer, M.A. and Mundy, N.I., 2008. Testing whether macroevolution follows microevolution: Are colour differences among swans (Cygnus) attributable to variation at the MC1R locus? *BMC Evolutionary Biology*, 8 (1), pp.249.

Poulton, E.B., 1890. Colours of animals. Kegan Paul, London.

Qanbar, R. and Bouvier, M., 2003. Role of palmitoylation/depalmitoylation reactions in G-protein-coupled receptor function. *Pharmacology and Therapeutics,* 97 (1), pp.1-33.

Rahmeh, R., Damian, M., Cottet, M., Orcel, H., Mendre, C., Durroux, T., Sharma, K.S., Durand, G., Pucci, B., Trinquet, E., Zwier, J.M., Deupi, X., Bron, P., Baneres, J.L., Mouillac, B. and Granier, S., 2012. Structural insights into biased G protein-coupled receptor signaling revealed by fluorescence spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America,* 109 (17), pp.6733-6738.

Raposo, G. and Marks, M.S., 2007. Melanosomes—dark organelles enlighten endosomal membrane transport. *Nature Reviews Molecular Cell Biology,* 8 (10), pp.786-797.

Ridley, A.J. and Hall, A., 1992. The small GTP-binding protein rho regulates the assembly of focal adhesions and actin stress fibers in response to growth factors. *Cell,* 70 (3), pp.389-399.

Ridley, A.J. and Hall, A., 1994. Signal transduction pathways regulating Rho-mediated stress fibre formation: requirement for a tyrosine kinase. *The EMBO Journal,* 13 (11), pp.2600-2610.

Rieder, S., Taourit, S., Mariat, D., Langlois, B. and Guérin, G., 2001. Mutations in the agouti (ASIP), the extension (MC1R), and the brown (TRP1) loci and their association to coat color phenotypes in horses (Equus caballus). *Mammalian Genome,* 12 (6), pp.450-455.

Ringholm, A., Klovins, J., Rudzish, R., Phillips, S., Rees, J.L. and Schiöth, H.B., 2004. Pharmacological characterization of loss of function mutations of the human melanocortin 1 receptor that are associated with red hair. *Journal of Investigative Dermatology,* 123 (5), pp.917-923.

Ritland, K., Newton, C. and Marshall, H.D., 2001. Inheritance and population structure of the white-phased "Kermode" black bear. *Current Biology,* 11 (18), pp.1468-1472.

Robbins, L.S., Nadeau, J.H., Johnson, K.R., Kelly, M.A., Roselli-Rehfuss, L., Baack, E., Mountjoy, K.G. and Cone, R.D., 1993. Pigmentation phenotypes of variant extension locus alleles result from point mutations that alter MSH receptor function. *Cell,* 72 (6), pp.827-834.

Roberts, D.W., Newton, R.A. and Sturm, R.A., 2007. MC1R expression in skin: is it confined to melanocytes? *Journal of Investigative Dermatology,* 127 (10), pp.2472-2473.

Rosenblum, E.B., Hoekstra, H.E. and Nachman, M.W., 2004. Adaptive reptile color variation and the evolution of the MCIR gene. *Evolution,* 58 (8), pp.1794-1808.

Rosenblum, E.B., Rompler, H., Schoneberg, T. and Hoekstra, H.E., 2010. Molecular and functional basis of phenotypic convergence in white lizards at White Sands. *Proceedings*

of the National Academy of Sciences of the United States of America, 107 (5), pp.2113-2117.

Roulin, A., 2004. The evolution, maintenance and adaptive function of genetic colour polymorphism in birds. *Biological Reviews*, 79 (4), pp.815-848.

Roulin, A. and Ducrest, A., 2011. Association between melanism, physiology and behaviour: a role for the melanocortin system. *European Journal of Pharmacology*, 660 (1), pp.226-233.

Rouzaud, F., Annereau, J.P., Valencia, J.C., Costin, G.E. and Hearing, V.J., 2003. Regulation of melanocortin 1 receptor expression at the mRNA and protein levels by its natural agonist and antagonist. *FASEB Journal*, 17 (14), pp.2154-2156.

Sánchez-Laorden, B.L., Sánchez-Más, J., Martínez-Alonso, E., Martínez-Menárguez, J.A., García-Borrón, J.C. and Jiménez-Cervantes, C., 2006. Dimerization of the human melanocortin 1 receptor: functional consequences and dominant-negative effects. *Journal of Investigative Dermatology*, 126 (1), pp.172-181.

Sanchez-Laorden, B.L., Jimenez-Cervantes, C. and Garcia-Borron, J.C., 2007. Regulation of human melanocortin 1 receptor signaling and trafficking by Thr-308 and Ser-316 and its alteration in variant alleles associated with red hair and skin cancer. *The Journal of Biological Chemistry*, 282 (5), pp.3241-3251.

Sanchez-Mas, J., Sanchez-Laorden, B., Guillo, L., Jimenez-Cervantes, C. and Garcia-Borron, J., 2005. The melanocortin-1 receptor carboxyl terminal pentapeptide is essential for MC1R function and expression on the cell surface. *Peptides*, 26 (10), pp.1848-1857.

Sánchez-Más, J., Gerritsen, I., García-Borrón, J. and Jiménez-Cervantes, C., 2004. Agonist-independent, high constitutive activity of the human mela-nocortin 1 receptor. *Experimental Dermatology*, 13 (9), pp.581-581.

Schneider, A., David, V.A., Johnson, W.E., O'Brien, S.J., Barsh, G.S., Menotti-Raymond, M. and Eizirik, E., 2012. How the Leopard Hides Its Spots: ASIP Mutations and Melanism in Wild Cats. *PLOS ONE*, 7 (12), pp.e50386.

Schülein, R., Hermosilla, R., Oksche, A., Dehe, M., Wiesner, B., Krause, G. and Rosenthal, W., 1998. A Dileucine Sequence and an Upstream Glutamate Residue in the Intracellular Carboxyl Terminus of the Vasopressin V2Receptor Are Essential for Cell Surface Transport in COS. M6 Cells. *Molecular Pharmacology*, 54 (3), pp.525-535.

Scott, G., Leopardi, S., Printup, S. and Madden, B.C., 2002. Filopodia are conduits for melanosome transfer to keratinocytes. *Journal of Cell Science*, 115 (7), pp.1441-1451.

Seamon, K.B., Padgett, W. and Daly, J.W., 1981. Forskolin: unique diterpene activator of adenylate cyclase in membranes and in intact cells. *Proceedings of the National Academy of Sciences of the United States of America*, 78 (6), pp.3363-3367.

Sharov, A.A., Fessing, M., Atoyan, R., Sharova, T.Y., Haskell-Luevano, C., Weiner, L., Funa, K., Brissette, J.L., Gilchrest, B.A. and Botchkarev, V.A., 2005. Bone morphogenetic protein (BMP) signaling controls hair pigmentation by means of cross-talk with the melanocortin receptor-1 pathway. *Proceedings of the National Academy of Sciences of the United States of America,* 102 (1), pp.93-98.

Shorten, M., 1954. *Squirrels.* Collins.

Shrake, A. and Rupley, J., 1973. Environment and exposure to solvent of protein atoms. Lysozyme and insulin. *Journal of Molecular Biology,* 79 (2), pp.351-371.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M. and Söding, J., 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology,* 7 (1).

Slominski, A. and Paus, R., 1993. Melanogenesis is coupled to murine anagen: toward new concepts for the role of melanocytes and the regulation of melanogenesis in hair growth. *Journal of Investigative Dermatology,* 101, pp. 90-97.

Slominski, A., Paus, R., Plonka, P., Chakraborty, A., Maurer, M., Pruski, D. and Lukiewicz, S., 1994. Melanogenesis during the anagen-catagen-telogen transformation of the murine hair cycle. *Journal of Investigative Dermatology,* 102 (6), pp.862-869.

Slominski, A., Plonka, P.M., Pisarchik, A., Smart, J.L., Tolle, V., Wortsman, J. and Low, M.J., 2005a. Preservation of eumelanin hair pigmentation in proopiomelanocortin-deficient mice on a nonagouti (a/a) genetic background. *Endocrinology,* 146 (3), pp.1245-1253.

Slominski, A., Wortsman, J., Plonka, P.M., Schallreuter, K.U., Paus, R. and Tobin, D.J., 2005b. Hair follicle pigmentation. *Journal of Investigative Dermatology,* 124 (1), pp.13-21.

Slominski, A., Tobin, D.J., Shibahara, S. and Wortsman, J., 2004. Melanin pigmentation in mammalian skin and its hormonal regulation. *Physiological Reviews,* 84 (4), pp.1155-1228.

Slominski, A., Wortsman, J., Luger, T., Paus, R. and Solomon, S., 2000. Corticotropin releasing hormone and proopiomelanocortin involvement in the cutaneous response to stress. *Physiological Reviews,* 80 (3), pp.979-1020.

Smith, C.C., 1981. The indivisible niche of Tamiasciurus: an example of nonpartitioning of resources. *Ecological Monographs,* pp.343-363.

Steiner, C.C., Weber, J.N. and Hoekstra, H.E., 2007. Adaptive variation in beach mice produced by two interacting pigmentation genes. *PLoS Biology,* 5 (9), pp.e219.

Steingrímsson, E., Copeland, N.G. and Jenkins, N.A., 2006. Mouse coat color mutations: from fancy mice to functional genomics. *Developmental Dynamics,* 235 (9), pp.2401-2411.

Strader, C.D., Fong, T.M., Tota, M.R., Underwood, D. and Dixon, R.A., 1994. Structure and function of G protein-coupled receptors. *Annual Review of Biochemistry,* 63 (1), pp.101-132.

Sturm, R.A., Teasdale, R.D. and Box, N.F., 2001. Human pigmentation genes: identification, structure and consequences of polymorphic variation. *Gene,* 277 (1), pp.49-62.

Suragani, R., Kamindla, R., Ehtesham, N. and Ramaiah, K., 2005. Interaction of recombinant human eIF2 subunits with eIF2B and eIF2α kinases. Biochem. Biophys. Res. Commun., 338 pp.1766–1772.

Swope, V.B., Jameson, J.A., McFarland, K.L., Supp, D.M., Miller, W.E., McGraw, D.W., Patel, M.A., Nix, M.A., Millhauser, G.L. and Babcock, G.F., 2012. Defining MC1R regulation in human melanocytes by its agonist α-melanocortin and antagonists agouti signaling protein and β-defensin 3. *Journal of Investigative Dermatology,* 132 (9), pp.2255-2262.

Takeuchi, S., Suzuki, H., Hirose, S., Yabuuchi, M., Sato, C., Yamamoto, H. and Takahashi, S., 1996. Molecular cloning and sequence analysis of the chick melanocortin 1-receptor gene. *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression,* 1306 (2), pp.122-126.

Takeuchi, S., Teshigawara, K. and Takahashi, S., 2000. Widespread expression of Agouti-related protein (AGRP) in the chicken: a possible involvement of AGRP in regulating peripheral melanocortin systems in the chicken. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research,* 1496 (2), pp.261-269.

Taussig, R. and Zimmermann, G., 1998. Type-specific regulation of mammalian adenylyl cyclases by G protein pathways. *Advances in Second Messenger and Phosphoprotein Research,* 32, pp.81-98.

Thayer, A.H., 1896. The law which underlies protective coloration. *The Auk,* 13 (2), pp.124-129.

Theron, E., Hawkins, K., Bermingham, E., Ricklefs, R.E. and Mundy, N.I., 2001. The molecular basis of an avian plumage polymorphism in the wild: a melanocortin-1-receptor point mutation is perfectly associated with the melanic plumage morph of the bananaquit, Coereba flaveola. *Current Biology,* 11 (8), pp.550-557.

Thevenet, P., Shen, Y., Maupetit, J., Guyon, F., Derreumaux, P. and Tuffery, P., 2012. PEP-FOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. *Nucleic Acids Research,* 40 (Web Server issue), pp.W288-93.

Thomsen, R. and Christensen, M.H., 2006. MolDock: a new technique for high-accuracy molecular docking. *Journal of Medicinal Chemistry,* 49 (11), pp.3315-3321.

Thörnwall, M., Dimitriou, A., Xu, X., Larsson, E. and Chhajlani, V., 1997. Immunohistochemical detection of the melanocortin 1 receptor in human testis, ovary and placenta using specific monoclonal antibody. *Hormone Research in Paediatrics,* 48 (5), pp.215-218.

Tobin, D. and Paus, R., 2001. Graying: gerontobiology of the hair follicle pigmentary unit. *Experimental Gerontology,* 36 (1), pp.29-54.

Tota, M., Smith, T., Mao, C., MacNeil, T., Mosley, R., Van der Ploeg, L. and Fong, T., 1999. Molecular interaction of agouti protein and agouti-related protein with human melanocortin receptors. *Biochemistry,* 38 (3), pp.897-904.

Trimming, P., 2011. *Grey Squirrel.* [online] Available at: <http://www.flickr.com/photos/peter-trimming/6583159839/> [Accessed 12 May 2014].

Tung, J.W., Heydari, K., Tirouvanziam, R., Sahaf, B., Parks, D.R., Herzenberg, L.A. and Herzenberg, L.A., 2007. Modern flow cytometry: a practical approach. *Clinics in Laboratory Medicine,* 27 (3), pp.453-468.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M. and Rozen, S.G., 2012. Primer3--new capabilities and interfaces. *Nucleic Acids Research,* 40 (15), pp.e115.

Våge, D.I., Fuglei, E., Snipstad, K., Beheim, J., Landsem, V.M. and Klungland, H., 2005. Two cysteine substitutions in the MC1R generate the blue variant of the arctic fox (Alopex lagopus) and prevent expression of the white winter coat. *Peptides,* 26 (10), pp.1814-1817.

Våge, D.I., Klungland, H., Lu, D. and Cone, R.D., 1999. Molecular and pharmacological characterization of dominant black coat color in sheep. *Mammalian Genome,* 10 (1), pp.39-43.

Våge, D.I., Lu, D., Klungland, H., Lien, S., Adalsteinsson, S. and Cone, R.D., 1997a. A non-epistatic interaction of agouti and extension in the fox, Vulpes vulpes. *Nature Genetics,* 15 (3), pp.311-315.

Våge, D.I., Lu, D., Klungland, H., Lien, S., Adalsteinsson, S. and Cone, R.D., 1997b. A non-epistatic interaction of agouti and extension in the fox, Vulpes vulpes. *Nature Genetics,* 15 (3), pp.311-315.

Våge, D., Nieminen, M., Anderson, D. and Røed, K., 2014. Two missense mutations in melanocortin 1 receptor (MC1R) are strongly associated with dark ventral coat color in reindeer (Rangifer tarandus). *Animal Genetics*.

Vassilatis, D.K., Hohmann, J.G., Zeng, H., Li, F., Ranchalis, J.E., Mortrud, M.T., Brown, A., Rodriguez, S.S., Weller, J.R., Wright, A.C., Bergmann, J.E. and Gaitanaris, G.A., 2003. The G protein-coupled receptor repertoires of human and mouse. *Proceedings of the National Academy of Sciences of the United States of America,* 100 (8), pp.4903-4908.

Venkatakrishnan, A., Deupi, X., Lebon, G., Tate, C.G., Schertler, G.F. and Babu, M.M., 2013. Molecular signatures of G-protein-coupled receptors. *Nature,* 494 (7436), pp.185-194.

Vidal, O., Araguas, R., Fernández, R., Heras, S., Sanz, N. and Pla, C., 2010. Melanism in guinea fowl (Numida meleagris) is associated with a deletion of Phenylalanine-256 in the MC1R gene. *Animal Genetics,* 41 (6), pp.656-658.

Vidal, O., Viñas, J. and Pla, C., 2010. Variability of the melanocortin 1 receptor (MC1R) gene explains the segregation of the bronze locus in turkey (Meleagris gallopavo). *Poultry Science,* 89 (8), pp.1599-1602.

Vignieri, S.N., Larson, J.G. and Hoekstra, H.E., 2010. The selective advantage of crypsis in mice. *Evolution,* 64 (7), pp.2153-2158.

Vrieling, H., Duhl, D.M., Millar, S.E., Miller, K.A. and Barsh, G.S., 1994. Differences in dorsal and ventral pigmentation result from regional expression of the mouse agouti gene. *Proceedings of the National Academy of Sciences of the United States of America,* 91 (12), pp.5667-5671.

Walker, W.P. and Gunn, T.M., 2010. Shades of meaning: the pigment-type switching system as a tool for discovery. *Pigment Cell and Melanoma Research,* 23 (4), pp.485-495.

Wallin, E. and von Heijne, G., 1995. Properties of N-terminal tails in G-protein coupled receptors: a statistical study. *Protein Engineering,* 8 (7), pp.693-698.

Walshaw, J. and Woolfson, D.N., 2001. Socket: a program for identifying and analysing coiled-coil motifs within protein structures. *Journal of Molecular Biology,* 307 (5), pp.1427-1450.

Wehrle-Haller, B., 2003. The Role of Kit-Ligand in Melanocyte Development and Epidermal Homeostasis. *Pigment Cell Research,* 16 (3), pp.287-296.

Wehrle-Haller, B. and Weston, J.A., 1995. Soluble and cell-bound forms of steel factor activity play distinct roles in melanocyte precursor dispersal and survival on the lateral neural crest migration pathway. *Development,* 121 (3), pp.731-742.

West, P.M. and Packer, C., 2002. Sexual selection, temperature, and the lion's mane. *Science,* 297 (5585), pp.1339-1343.

Whippey, P., 2012. *Melanic grey squirrel*. [online] Available at: <http://www.flickr.com/photos/peter-trimming/6583159839/> [Accessed 12 May 2014].

Whiteley, D.A., Owen, D.F. and Smith, D.A., 1997. Massive polymorphism and natural selection in Donacilla cornea (Poli, 1791) (Bivalvia: Mesodesmatidae). *Biological Journal of the Linnean Society,* 62 (4), pp.475-494.

Wilson, B.D., Ollmann, M.M., Kang, L., Stoffel, M., Bell, G.I. and Barsh, G.S., 1995. Structure and function of ASP, the human homolog of the mouse agouti gene. *Human Molecular Genetics,* 4 (2), pp.223-230.

Wlasiuk, G. and Nachman, M.W., 2007. The genetics of adaptive coat color in gophers: coding variation at Mc1r is not responsible for dorsal color differences. *Journal of Heredity,* 98 (6), pp.567-574.

Xiang, Z., 2006. Advances in homology protein structure modeling. *Current Protein and Peptide Science,* 7 (3), pp.217-227.

Yang, Y., 2011. Structure, function and regulation of the melanocortin receptors. *European Journal of Pharmacology,* 660 (1), pp.125-130.

Yang, Y., Chen, M., Lai, Y., Gantz, I., Georgeson, K.E. and Harmon, C.M., 2002. Molecular determinants of human melanocortin-4 receptor responsible for antagonist SHU9119 selective activity. *The Journal of Biological Chemistry,* 277 (23), pp.20328-20335.

Yang, Y., Dickinson, C., Haskell-Luevano, C. and Gantz, I., 1997. Molecular basis for the interaction of [Nle4, D-Phe7] melanocyte stimulating hormone with the human melanocortin-1 receptor. *The Journal of Biological Chemistry,* 272 (37), pp.23000-23010.

Yao, X., Parnot, C., Deupi, X., Ratnala, V.R., Swaminath, G., Farrens, D. and Kobilka, B., 2006. Coupling ligand structure to specific conformational switches in the β2-adrenoceptor. *Nature Chemical Biology,* 2 (8), pp.417-422.

Yu, W., Wang, C., Xin, Q., Li, S., Feng, Y., Peng, X. and Gong, Y., 2013. Non-synonymous SNPs in MC1R gene are associated with the extended black variant in domestic ducks (Anas platyrhynchos). *Animal Genetics,* 44 (2), pp.214-216.

Zhang, Y., Arakaki, A.K. and Skolnick, J., 2005. TASSER: an automated method for the prediction of protein tertiary structures in CASP6. *Proteins: Structure, Function, and Bioinformatics,* 61 (S7), pp.91-98.

Appendices

Appendix 1. Common protocols and buffers

Medium for bacterial growth (1L)

16 g of Luria-Bertani (LB) powder (Sigma) was dissolved in 950 ml of deionised water. The medium was autoclaved for 20 minutes at 15 psi. The medium was allowed to cool to 55ºC and 50µg/ml of ampicillin was added. Medium was cooled to 37 ºC before use.

Agar plates for bacterial growth

32 g of LB agar (Sigma) was dissolved in 950 ml of deionised water. The agar was autoclaved for 20 minutes at 15 psi. The medium was allowed to cool to 55ºC and 50µg/ml of ampicillin was added. The agar was poured into petri dishes using the aseptic technique. Plates were left to set and stored at 4ºC.

Transfer buffer salts: 1 Litre: 10 ×

30.3 g Tris base
144 g glycine
Made up to 1L with ddH_2O

Transfer buffer: 1 Litre: 1 ×

100 ml 10 × Transfer buffer salts

200 ml methanol

700 ml ddH_2O

TBS: 1 Litre: 10 ×

87.7 g $NaCl_2$

12.1 g Tris base

4 ml HCl to pH 8

Made up to 1L with ddH_2O

TBS-Tween: 1 Litre: 1 ×

100 ml 10 × TBS

10 ml Tween 20

890 ddH_2O

Blocking buffer: 100ml

5 g BSA

Made up to 100ml with TBS-Tween

TBE: 1 Litre: 10 ×

108 g Tris base

55 g boric acid

7.5 g EDTA

Made up to 1L with ddH_2O

Appendix 2. Primers

Primers used to amplify genomic DNA

Primer Name	Sequence
MSHR4F	5′-TGC TTC CTG GAC AGG ACT ATG-3′
MC1RinvR2	5′-ATG GAG ATG TAG CGG TCC AC -3′
MC1Rdel	5′- AAC GCA CTG GAG ACG ACC ATC-3′
MC1RBF	5′-CTG GTG AGC ACC TTC CTA CTG-3′
MC1RBR	5′-CCA GCA GTA GGA AGG TG-3′
MC1RGR	5'-GTC TCC AGT GCG TTG CT-3'
MSHF1	5' -CTG CAA GCC CGA CCT CTC-3'
MSHR9	5'-TCT TGA TGA TGG CGT TTT TG-3'
MSHR4	5'-CCG TAG CGC TTG TCC TTG-3'
MC1Rer6	5'-CTG GGC TGG AGA CCA GAA TC-3'
ASP2_13LF1	5'-AGT ACT CCG CCC TCT GGA TA-3'
ASP3sqR11	5'-TCT GCT TCT TTT CTG CTG AT-3'
Sqintron2to3F1	5'-CCC TCT GCT CCT TCC ATT TT-3'
Aspintron3to4R1	5'-AAT GAG AAC TCC CAG GCC TAC-3'
ASP3to4sqF11	5'-ATG GAC AGC TCC CGC ATT T-3'
ASP4sqrex30	5'-AGG AAG CTT TGA GTG GAC GA-3'
MC1RexpF1	5'-CAC C**AT G**CC TGT ACA GAG GAG GCT CC-3'
MC1RexpR1	5'-CCC AGG AGC ACA GCA GCA CCT CC-3'

Primers for site-directed mutagenesis

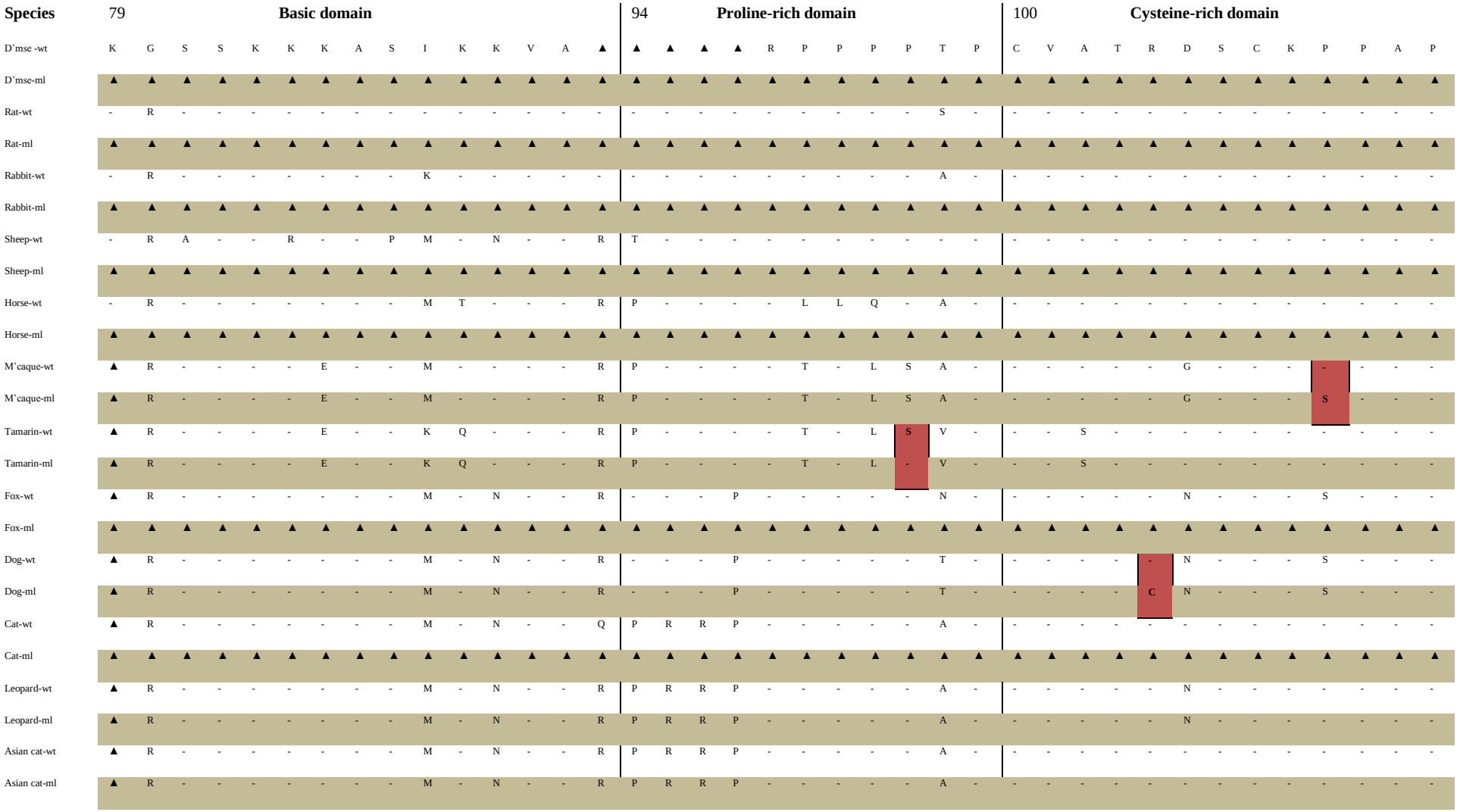
Primer name	Sequence
GGFPN88AF	5'-GAC CTG CTG GTG AGC ACC AGC GCC GCA CTG GAG ACG ACC ATC TTC C-3'
GGFPN88AR	5'-GGA AGA TGG TCG TCT CCA GTG CGG CGC TGG TGC TCA CCA GCA GGT C-3'
GGFPE91KF	5'-GAC CTG CTG GTG AGC ACC AGC AAC GCA CTG AAG ACG ACC ATC TTC C-3'
GGFPE91KR	5'-G GAA GAT GGT CGT CTT CAG TGC GTT GCT GGT GCT CAC CAG CAG GTC-3'
GGFPdel4F	5'-TCG GAC CTG CTG GTG AGC ACC GAG ACG ACC ATC TTC-3'
GGFPdelR	5'-GAA GAT GGT CGT CTC GGT GCT CAC CAG CAG GTC CGA-3'
GGFPdel2F	5'-TCG GAC CTG CTG GTG AGC ACC GCA CTG GAG ACG ACC ATC-3'
GGFPdel2R	5'-GAT GGT CGT CTC CAG TGC GGT GCT CAC CAG CAG GTC CGA-3'
BGFPE91KF	5'-AGC ACC TTC CTA CTG CTG AAG GTG GGT GCC CTG GCA ACG CCA G-3'
BGFPE91KR	5'-CTG GCG TTG CCA GGG CAC CCA CCT TCA GCA GTA GGA GGT GCT-3'

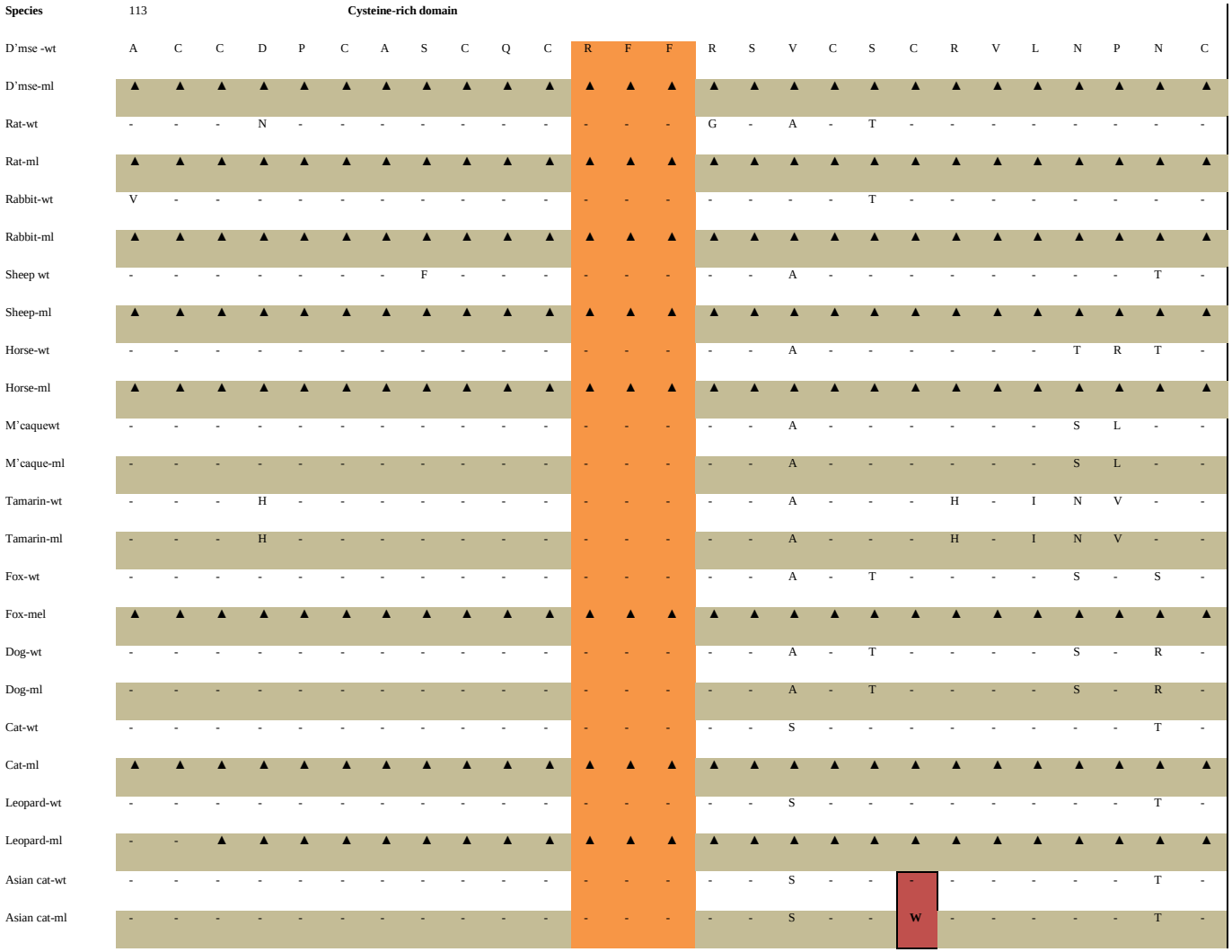
Appendix 3. Amino acid sequence alignment of agouti signalling protein variants associated with melanism

ASIP sequences from wildtype (wt) are shown in contrast to the melanic (mel) variant underneath for the deer mouse ((Kingsley et al., 2009) (Peromyscus maniculatus) (Accession: ACV71928.1), rat (Rattus norvegicus) (Kuramoto et al., 2001) (Accession: AB045587), rabbit (Oryctolagus cuniculus) (Accession: AM748788), sheep (Ovis aries) (Accession: EU057181), horse (Equus caballus) (Accession: AF288358.1), Macaque (Macaca nigra) (Accession: EF094485), Golden lion tamarin (Leontopithecus rosalia) (Mundy and Kelly, 2006) (Accession: EF094499), fox (Vulpes Vulpes) (Våge et al., 1997a), dog (Canus lupus) (Kerns et al., 2004), Leopard (Panthera pardus) (Schneider et al., 2012), Asian Golden cat (Pardofelis temminckii) (Schneider et al., 2012). Numbers are according to the deer mouse sequence. Domains of the protein are indicated. The boxed areas highlighted red show amino acid substitutions associated with melanism. The orange boxed area highlights the highly conserved RFF motif. (D'mse= deer mouse, M'caque= Macaque, Tamarin= Golden lion tamarin, (the melanic variant is called Black lion tamarin), Asian cat= Asian golden cat.)

Species	1																						24																	
	Signal peptide																						**N-Terminus**																	
D'mse -wt	M	D	V	T	R	L	L	L	A	T	L	V	G	F	L	C	F	L	A	V	Y	S	H	L	V	P	E	E	T	L	R	D	D	K	S	L	K	T	N
D'mse-ml	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
Rat-wt	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	H	-	-	-	-	F	-	-	-	-	G	-	-	R	-	-	-	S	-
Rat-ml	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	H	-	-	-	-	F	-	-	-	-	G	-	-	R	-	-	▲	▲	▲
Rabbit-wt	-	N	-	-	-	-	-	Q	-	-	-	L	V	-	-	-	-	-	T	A	-	-	-	-	A	-	-	-	-	P	T	-	-	Q	-	-	R	S	-
Rabbit-ml	-	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
Sheep-wt	-	-	-	S	-	-	F	-	-	-	-	L	V	C	-	-	-	-	S	A	-	-	-	-	A	-	-	-	K	P	-	-	E	R	N	-	-	N	-
Sheep-ml	-	-	-	S	-	-	F	-	-	-	-	L	V	C	-	-	-	-	S	A	-	-	-	-	A	-	-	-	K	P	-	-	▲	▲	▲	▲	▲	▲	▲
Horse-wt	-	-	-	I	H	-	F	-	-	-	-	L	V	S	-	-	-	-	T	A	-	-	-	-	S	-	-	-	K	P	K	-	-	R	-	-	R	N	-
Horse-ml	-	-	-	I	H	-	F	-	-	-	-	L	V	S	-	-	-	-	T	A	-	-	-	-	S	-	-	-	K	P	K	-	-	R	-	-	R	N	-
M'caque-wt	-	-	-	-	-	-	-	-	-	-	-	L	V	-	-	-	-	F	T	A	-	-	-	P	P	-	-	-	K	-	-	-	-	R	-	-	R	S	-
M'caque-ml	-	-	-	-	-	-	-	-	-	-	-	L	V	-	-	-	-	F	T	A	-	-	-	P	P	-	-	-	K	-	-	-	-	R	-	-	R	S	-
Tamarin-wt	-	-	-	-	-	-	-	-	-	-	-	L	V	-	-	-	C	F	-	A	-	-	-	-	P	-	-	-	K	-	-	-	-	R	-	-	R	S	-
Tamarin-ml	-	-	-	-	-	-	-	-	-	-	-	L	V	-	-	-	C	F	-	A	-	-	-	-	P	-	-	-	K	-	-	-	-	R	-	-	R	S	-
Fox-wt	-	N	I	F	-	-	-	-	-	-	-	L	V	S	-	-	-	-	T	A	-	-	-	-	A	▲	-	-	K	P	K	-	-	R	-	-	R	S	-
Fox-ml	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
Dog-wt	-	N	I	F	-	-	-	-	-	-	-	L	V	S	-	-	-	-	T	A	-	-	-	-	A	▲	-	-	K	P	K	-	-	R	-	-	R	S	-
Dog-ml	-	N	I	F	-	-	-	-	-	-	-	L	V	S	-	-	-	-	T	A	-	-	-	-	A	▲	-	-	K	P	K	-	-	R	-	-	R	S	-
Cat-wt	-	N	I	L	-	-	-	-	-	-	-	L	V	C	-	-	L	-	T	A	-	-	-	-	A	-	-	-	K	P	-	-	-	R	N	-	R	S	-
Cat-ml	-	N	I	L	-	-	-	-	-	-	-	L	V	C	-	-	L	-	T	A	-	-	-	-	A	-	-	-	K	P	-	-	-	R	N	-	R	S	-
Leopard-wt	-	N	I	L	-	-	-	-	-	-	-	L	V	S	-	-	L	-	T	A	-	-	-	-	A	-	-	-	K	P	-	-	-	R	N	M	R	S	-
Leopard-ml	-	N	I	L	-	-	-	-	-	-	-	L	V	S	-	-	L	-	T	A	-	-	-	-	A	-	-	-	K	P	-	-	-	R	N	M	R	S	-
Asian cat-wt	-	N	I	L	-	-	-	-	-	-	-	L	V	S	-	-	L	-	T	A	-	-	-	-	A	-	-	-	K	P	-	-	-	R	N	-	R	S	-
Asian cat-ml	-	N	I	L	-	-	-	-	-	-	-	L	V	S	-	-	L	-	T	A	-	-	-	-	A	-	-	-	K	P	-	-	-	R	N	-	R	S	-

Species	40																	57																				
	N-Terminus																	**Basic domain**																				
D'mse -wt	S	S	T	D	C	K	D	F	S	▲	V	S	I	V	A	L	K	K	K	S	K	K	I	S	I	Q	E	A	E	K	Q	K	R	L	E	A	E	K	R
D'mse-ml	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
Rat-wt	-	-	I	N	S	L	-	-	-	S	-	-	-	-	-	-	N	-	-	-	-	-	-	-	R	K	-	-	▲	▲	▲	▲	▲	▲	▲	▲	-	-	-
Rat-ml	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
Rabbit-wt	-	-	-	N	L	L	E	-	-	S	-	-	-	-	-	-	-	-	-	-	-	D	-	-	-	K	-	-	▲	▲	▲	▲	▲	▲	▲	▲	-	-	K
Rabbit-ml	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
Sheep-wt	-	-	M	N	L	L	-	-	P	S	-	-	-	-	-	-	N	-	-	-	-	-	-	-	R	N	-	-	▲	▲	▲	▲	▲	▲	▲	▲	-	-	K
Sheep-ml	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
Horse-wt	-	-	M	N	L	L	-	S	P	S	-	-	-	M	-	-	N	-	-	-	-	-	-	-	R	K	-	-	▲	▲	▲	▲	▲	▲	▲	▲	-	-	K
Horse-ml	-	-	M	N	L	L	-	S	P	S	-	-	-	M	-	-	N	-	-	-	-	-	-	-	R	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
M'caque-wt	-	-	V	N	L	L	-	-	P	S	-	-	-	-	-	-	N	-	N	-	-	Q	-	-	R	K	-	-	▲	▲	▲	▲	▲	▲	▲	▲	-	-	K
M'caque-ml	-	-	V	N	L	L	-	-	P	S	-	-	-	-	-	-	N	-	N	-	-	Q	-	-	R	K	-	-	▲	▲	▲	▲	▲	▲	▲	▲	-	-	K
Tamarin-wt	-	-	V	N	L	L	-	L	P	S	-	-	-	-	-	-	N	-	-	-	-	-	-	-	R	K	-	-	▲	▲	▲	▲	▲	▲	▲	▲	-	N	K
Tamarin-ml	-	-	V	N	L	L	-	L	P	S	-	-	-	-	-	-	N	-	-	-	-	-	-	-	R	K	-	-	▲	▲	▲	▲	▲	▲	▲	▲	-	N	K
Fox-wt	-	-	V	N	L	L	-	-	P	S	-	-	-	-	-	-	N	-	-	-	-	-	-	-	R	K	-	-	▲	▲	▲	▲	▲	▲	▲	▲	-	-	K
Fox-ml	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
Dog-wt	-	-	V	N	L	L	-	-	P	S	-	-	-	-	-	-	N	-	-	-	-	-	-	-	R	K	-	-	▲	▲	▲	▲	▲	▲	▲	▲	-	-	K
Dog-ml	-	-	V	N	L	L	-	-	P	S	-	-	-	-	-	-	N	-	-	-	-	-	-	-	R	K	-	-	▲	▲	▲	▲	▲	▲	▲	▲	-	-	K
Cat-wt	-	-	M	N	M	L	-	L	-	S	-	-	-	-	-	-	N	-	-	-	-	-	-	-	R	K	-	-	▲	▲	▲	▲	▲	▲	▲	▲	-	-	K
Cat-ml	-	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
Leopard-wt	-	-	M	N	L	L	-	L	P	S	-	-	-	-	-	-	N	-	-	-	-	-	-	-	R	K	-	-	▲	▲	▲	▲	▲	▲	▲	▲	-	-	K
Leopard-ml	-	-	M	N	L	L	-	L	P	S	-	-	-	-	-	-	N	-	-	-	-	-	-	-	R	K	-	-	▲	▲	▲	▲	▲	▲	▲	▲	-	-	K
Asian cat-wt	-	-	M	N	L	L	-	L	P	S	-	-	-	-	-	-	N	-	-	-	-	-	-	-	R	K	-	-	▲	▲	▲	▲	▲	▲	▲	▲	-	-	K
Asian cat-ml	-	-	M	N	L	L	-	L	P	S	-	-	-	-	-	-	N	-	-	-	-	-	-	-	R	K	-	-	▲	▲	▲	▲	▲	▲	▲	▲	-	-	K


Species
79
Basic domain
94
Proline-rich domain
100
Cysteine-rich domain
D'mse -wt K G S S K K K A S I K K V A ▲ ▲ ▲ ▲ ▲ R P P P P T P C V A T R D S C K P P A P
D'mse-ml
Rat-wt - R - S - - - - - - - - - - - - -
Rat-ml
Rabbit-wt - R - - - - - - - K - - - - - - - - - - - - - - A - - - - - - - - - - - - -
Rabbit-ml
Sheep-wt - R A - - R - - P M - N - - R T -
Sheep-ml
Horse-wt - R - - - - - - - M T - - - R P - - - - L L Q - A - - - - - - - - - - - - -
Horse-ml
M'caque-wt ▲ R - - - - E - - M - - - - R P - - - - T - L S A - - - - - - G - - - - - - -
M'caque-ml ▲ R - - - - E - - M - - - - R P - - - - T - L S A - - - - - - G - - - S - - -
Tamarin-wt ▲ R - - - - E - - K Q - - - R P - - - - T - L S V - - - S - - - - - - - - - -
Tamarin-ml ▲ R - - - - E - - K Q - - - R P - - - - T - L - V - - - S - - - - - - - - - -
Fox-wt ▲ R - - - - - - - M - N - - R - - - P - - - - - N - - - - - - N - - - S - - -
Fox-ml
Dog-wt ▲ R - - - - - - - M - N - - R - - - P - - - - - T - - - - - - - - - - - - - -
Dog-ml ▲ R - - - - - - - M - N - - R - - - P - - - - - T - - - - - - - - - - - - - -
Cat-wt ▲ R - - - - - - - M - N - - Q P R R P - - - - - A - - - - - - - - - - - - - -
Cat-ml
Leopard-wt ▲ R - - - - - - - M - N - - R P R R P - - - - - A - - - - - - N - - - - - - -
Leopard-ml ▲ R - - - - - - - M - N - - R P R R P - - - - - A - - - - - - N - - - - - - -
Asian cat-wt ▲ R - - - - - - - M - N - - R P R R P - - - - - A - - - - - - - - - - - - -
Asian cat-ml ▲ R - - - - - - - M - N - - R P R R P - - - - - A - - - - - - - - - - - - -


Species
113
Cysteine-rich domain
D'mse -wt A C C D P C A S C Q C R F F R S V C S C R V L N P N C
D'mse-ml
Rat-wt - - - N - - - - - - - - - - G - A - T - - - - - - - -
Rat-ml
Rabbit-wt V - - - - - - - - - - - - - - - - T - - - - - - - -
Rabbit-ml
Sheep wt - - - - - - - F - - - - - - - - A - - - - - - - T -
Sheep-ml
Horse-wt - - - - - - - - - - - - - - - - A - - - - - T R T -
Horse-ml
M'caquewt - - - - - - - - - - - - - - - - A - - - - - S L - -
M'caque-ml - - - - - - - - - - - - - - - - A - - - - - S L - -
Tamarin-wt - - - H - - - - - - - - - - - - A - - - H - I N V - -
Tamarin-ml - - - H - - - - - - - - - - - - A - - - H - I N V - -
Fox-wt - - - - - - - - - - - - - - - - A - T - - - - S - S -
Fox-mel
Dog-wt - - - - - - - - - - - - - - - - A - T - - - - S - R -
Dog-ml - - - - - - - - - - - - - - - - A - T - - - - S - R -
Cat-wt - - - - - - - - - - - - - - - - S - - - - - - - T -
Cat-ml
Leopard-wt - - - - - - - - - - - - - - - - S - - - - - - - T -
Leopard-ml - - ▲
Asian cat-wt - - - - - - - - - - - - - - - - S - - - - - - - T -
Asian cat-ml - - - - - - - - - - - - - - - - S - - W - - - - - T -

Appendix 4. Predicted "knobs into holes" molecular interactions of the melanocortin-1 receptor

"Knobs into holes" interactions occur where hydrophobic amino acids from one helix become interlocked into hydrophobic holes created by amino acids on an adjacent helix. Letters refer to amino acids and numbers refer to the amino acid numbers. TM= transmembrane helix.

Knobs		Holes	
TM1	L 45	TM7	A 282
			I 285
			C 286
			I 289
TM1	V 57	TM2	C 75
			A 78
			L79
			L 82
TM1	A 59	TM8	L 302
			T 305
			L 306
			V 309
TM2	C 72	TM4	L 155
			A 158
			R 159
			I 162
TM2	F 73	TM3	F 131
			A 134
			I 135
			D 138
TM2	S 80	TM3	S 124
			S 127
			S 128
			F 131
TM2	L 82	TM1	L 50
			N 53
			M 54
			V 57

Knobs		Holes	
TM3	F 131	TM2	F 73
			C 76
			L 77
			S 80
TM3	A 136	TM5	M 200
			L 203
			Y 204
			M 207
TM4	I 165	TM3	S 127
			C 130
			F 131
			A 134
TM5	Y 204	TM6	A 237
			L 240
			T 241
			L 244
TM6	L 244	TM5	L 197
			M 200
			A 201
			Y 204
TM6	F 247	TM3	M 125
			S 128
			L 129
			L 132
TM6	C 250	TM7	L 283
			C 286
			N 287
			V 290

Knobs		Holes	
TM7	I 289	TM1	L 45
			V 48
			S 49
			E 52
TM7	I 294	TM6	T 239
			T 242
			L 243
			V 246
TM8	V 309	TM1	L 55
			V 58
			A 59
			K 62

Appendix 5. Residues predicted to be in the binding pocket of the MC1R-wt and MC1RΔ24 receptors

Numbers in the residue column are MC1R-wt numbering with mutant numbering in brackets. A cross refers to a single piece of evidence that the residue in question is involved in binding and/or activation in the binding pocket of the MC1R. Data were obtained from three sources; functional studies, crystallography and computer predictions. Functional studies are further divided into data from MC1R and MC4R experiments. Computer predictions are divided into data using the MC1R-wt and MC1RΔ24 (Phyre2 model) with Fpocket, MC1R-wt and MC1RΔ24 (I-TASSER model) with Fpocket and computer predictions were also made using MC1R-wt (Phyre2 model) with Pepsite. (No data were available using the MC1RΔ24 for Pepsite.) Numbers refer to a scoring system: 7 points are given for a residue which is both MC1R functional data and conserved with the MC1R-wt/MC1RΔ24. 6 points are given for a residue which is found in MC1R functional data but not conserved. 5 points are given for a residue which is found in MC4R functional data and is conserved. 4 points are given for a residue which is found in MC4R functional data but is not conserved. 3 points are given for a residue which is found in crystallography data which is conserved. 2 points are given for a residue which is found in crystallography data but is not conserved. 1 point is given for a residue identified by computer prediction. Total scores exceeding 15 points are highlighted green. Green crosses are conserved and black crosses are not conserved. WT= MC1R-wt, Δ24 = MC1RΔ24, A2AR = adenosine. (Robbins et al., 1993; Våge et al., 1997a; Yang et al., 1997; Lu, Vage and Cone, 1998; Haskell-Luevano et al., 2001; Más et al., 2002; Yang et al., 2002; Ringholm et al., 2004; Beaumont et al., 2005; Chai et al., 2005; Benned-Jensen, Mokrosinski and Rosenkilde, 2011; Doré et al., 2011; Lebon et al., 2011)

	Experimental data			Computer predictions					
	Functional data		Cry'phy data						
Residue	MC1R	MC4R	A2AR	WT Phyre Fpocket	Δ24 Phyre Fpocket	WT Phyre Pepsite	WT TASSER Fpocket	Δ24 TASSER Fpocket	Tot.
Score	7 or 6	5 or 4	3 or 2	1	1	1	1	1	
80S (80)				+	+			+	3
81D (81)	+			+	+				9
83L (83)								+	1
84V (84)				+	+			+	3
88N (▲)				+					1
89A (▲)	++			+					13
90L (▲)	+								7
91E (▲)	+++++	++			+				46
92T (▲)				+			+		2
95F (87)				+			+	+	3
96L (88)				+	+		+	+	4
98L (90)	+			+					8
99E (91)				+		++	+		4
100V (92)				+	+		+		3
103L (95)					+				1
104A (96)					+				1
106P (98)					+				1
107A (99)					+				1
114D (106)	++	++		+	+				26
116V (108)								+	1
117I (109)					+				1
118D (110)	+++	++		+	+	+	+		35
119V (111)					+				1
120L (112)								+	1

Residue	MC1R	MC4R	A2AR	WT Phyre Fpocket	Δ24 Phyre Fpocket	WT Phyre Pepsite	WT TASSER Fpocket	Δ24 TASSER Fpocket	Tot.
121T (113)	+	+	+	+	+		+		17
122C (114)	++		+	+	+		+		19
123G (115)								+	1
124S (116)				+	+			+	3
125M (117)		+	+	+	+				8
126V (118)					+				1
127S (119)								+	1
128S (120)				+	+				2
166W (158)								+	1
169S (161)								+	1
170I (162)								+	1
172S (164)	+	+			+				12
173S (165)					+				1
176F (168)		+			+				5
180Y (172)	+	+							11
181N (173)	+								6
182H (174)	+								7
184A (176)				+	+				2
185V (177)				+	+				2
186L (178)					+				1
187L (179)					+		+		2
188C (180)			++		+		+		6
189L (181)			+	+	+		+		5
190V (182)				+	+				2
192F (184)	+						+		8
193F (185)	+			+					8
194L (186)				+					1
197L (189)			++	+			+		8
201L (193)			+						2
247F (239)				+	+				2
248F (240)				+					

Residue	MC1R	MC4R	A2AR	WT Phyre Fpocket	Δ24 Phyre Fpocket	WT Phyre Pepsite	WT TASSER Fpocket	Δ24 TASSER Fpocket	Tot.
251W (243)	+	+	+	+	+	+	+		19
252G (244)				+					1
254F (246)	++	+++	++	+	+	++	+		38
255F (247)	+	+	+	+					15
257H (249)		+				++			7
258L (250)		+	++		+				10
260L (252)		+							4
261I (253)					+		+		2
267H (259)							+		1
268P (260)							+		1
269A (261)							+		1
270C (262)							+		1
273V (265)							+		1
274F (266)	+			+	+		+		10
277F (273)	+	+	++	+	+	++	+		20
278N (270)				+	+		+		3
281L (273)	+		++	+	+		+		13
284I (276)			+	+	+	+	+		6
285I (277)			+	+	+				4
287N (279)				+	+				2
288S (280)				+	+				2

Appendix 6.

McRobie, H.R., Thomas, A. and Kelly, J., 2009. The genetic basis of melanism in the gray squirrel (*Sciurus carolinensis*). *The Journal of Heredity*, 100 (6), pp.709-714.

Appendix 7.

McRobie, H.R., King, L.M., Fanutti, C., Symmons, M.F. and Coussons, P.J., 2014. Agouti signalling protein is an inverse agonist to the wildtype and agonist to the melanic variant of the melanocortin-1 receptor in the grey squirrel (*Sciurus carolinensis*). *FEBS Letters.588 (14) pp.2335-2343.*

Appendix 8.

McRobie, H.R., King, L.M., Fanutti, C., Coussons, P.J., Moncrief, N.D. and Thomas, A.P., 2014. Melanocortin-1 receptor (MC1R) gene sequence variation and melanism in the gray (*Sciurus carolinensis*), fox (*Sciurus niger*), and red (*Sciurus vulgaris*) squirrel. *The Journal of Heredity*, 105 (3), pp.423-428.