

**Molekularbiologische Charakterisierung der  
beiden Untereinheiten des Mega-Hämocyanins  
der Schnecke *Melanoides tuberculata***

Dissertation  
zur Erlangung des Grades  
**„Doktor der Naturwissenschaften“**

am Fachbereich Biologie  
der Johannes Gutenberg-Universität  
in Mainz

Meik Neufurth  
geboren in Hachenburg

Mainz, 2011

Dekan:

1. Berichterstatter:

2. Berichterstatter:

Tag der mündlichen Prüfung: 16.12.2011

*D77 (Dissertation Universität Mainz)*



*Für meine Omi*



# Inhaltsverzeichnis

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Einleitung</b> .....                              | <b>1</b>  |
| 1.1      | Höheres Leben benötigt Sauerstoff.....               | 1         |
| 1.2      | Respiratorische Proteine.....                        | 1         |
| 1.2.1    | Mollusken-Hämocyanine: Eine Übersicht .....          | 3         |
| 1.2.2    | Das Hämocyanin der Gastropoden .....                 | 6         |
| 1.2.3    | Die Genstruktur der Mollusken-Hämocyanine.....       | 8         |
| 1.3      | Das Mega-Hämocyanin der Cerithioidea .....           | 11        |
| 1.4      | Die Cerithioidea .....                               | 13        |
| 1.5      | Ziel der Arbeit .....                                | 15        |
| <b>2</b> | <b>Material und Methoden</b> .....                   | <b>16</b> |
| 2.1      | Tiere .....  | 16        |
| 2.1.1    | <i>Melanoides tuberculata</i> .....                  | 16        |
| 2.2      | Chemikalien und Geräte .....                         | 17        |
| 2.3      | Allgemeine Vorsichtsmaßnahmen.....                   | 17        |
| 2.4      | Mikrobiologische Methoden .....                      | 18        |
| 2.4.1    | Nährmedien und Agarplatten .....                     | 18        |
| 2.4.2    | Plattenkultur .....                                  | 19        |
| 2.4.3    | Flüssigkultur.....                                   | 19        |
| 2.4.4    | Glyzerinkultur .....                                 | 19        |
| 2.4.5    | Bakterienstämme und Vektoren.....                    | 20        |
| 2.4.5.1  | Bakterienstämme:.....                                | 20        |
| 2.4.5.2  | Vektoren .....                                       | 20        |
| 2.5      | Molekularbiologische Methoden .....                  | 20        |
| 2.5.1    | Präparation von Nukleinsäuren.....                   | 20        |
| 2.5.1.1  | Präparation genomischer DNA .....                    | 21        |
| 2.5.1.2  | Präparation genomischer DNA mit CTAB.....            | 21        |
| 2.5.1.3  | Präparation von Gesamt-RNA .....                     | 22        |
| 2.5.1.4  | Isolierung von mRNA aus Gesamt-RNA .....             | 22        |
| 2.5.2    | RNA-Aufreinigung mit rDNase-Verdau .....             | 23        |
| 2.5.3    | Phenol/Chloroform Extraktion von Nukleinsäuren ..... | 23        |
| 2.5.4    | Ethanolpräzipitation von Nukleinsäuren .....         | 23        |
| 2.5.5    | Photometrische Konzentrationsbestimmung.....         | 24        |
| 2.5.6    | Reverse Transkription (RT) .....                     | 24        |
| 2.5.7    | PCR Techniken .....                                  | 25        |
| 2.5.7.1  | Primer.....  | 25        |
| 2.5.7.2  | PCR.....   | 25        |
| 2.5.7.3  | Nested PCR .....                                     | 27        |
| 2.5.7.4  | Touchdown PCR.....                                   | 28        |

|          |   |           |
|----------|---|-----------|
| 2.5.7.5  | Band-stab PCR.....  | 28        |
| 2.5.7.6  | Klon PCR.....   | 29        |
| 2.5.7.7  | GeneRacer .....   | 30        |
| 2.5.7.8  | PCR-Aufreinigung .....  | 31        |
| 2.5.8    | Gelelektrophoretische Techniken .....   | 31        |
| 2.5.8.1  | Agarose-Gelelektrophorese von DNA.....  | 31        |
| 2.5.8.2  | Agarose-Gelelektrophorese von RNA.....  | 33        |
| 2.5.8.3  | Extraktion von DNA aus Agarosegelen .....                                     | 35        |
| 2.5.9    | Klonierung von PCR-Produkten.....   | 35        |
| 2.5.9.1  | Blau-Weiß-Selektion .....   | 38        |
| 2.5.9.2  | Selektion mit Hilfe des letalen <i>E. coli</i> Gens.....                      | 38        |
| 2.5.9.3  | Präparation von Plasmid-DNA.....  | 39        |
| 2.5.9.4  | Restriktionsverdau von Plasmid-DNA .....                                      | 39        |
| 2.6      | Sequenzanalyse .....  | 40        |
| 2.6.1    | DNA-Sequenzierung.....  | 40        |
| 2.6.2    | Sequenzanalyse .....  | 40        |
| 2.7      | Phylogenie .....  | 41        |
| <b>3</b> | <b>Ergebnisse.....</b>  | <b>42</b> |
| 3.1      | Ermittlung erster Sequenzdaten .....  | 42        |
| 3.2      | Sequenzierung der MtH <sub>400</sub> -Untereinheit .....                      | 45        |
| 3.2.1    | Die 5'-UTR der MtH <sub>400</sub> -cDNA.....                                  | 47        |
| 3.2.2    | Die 3'-UTR der MtH <sub>400</sub> -cDNA.....                                  | 47        |
| 3.2.3    | Die MtH <sub>400</sub> -cDNA besitzt ein Signalpeptid .....                   | 48        |
| 3.2.4    | Die MtH <sub>400</sub> -cDNA besitzt eine „Translation initiation site“ ..... | 49        |
| 3.3      | Sequenzierung der MtH <sub>550</sub> -Untereinheit .....                      | 50        |
| 3.3.1    | Die 5'-UTR der MtH <sub>550</sub> -cDNA.....                                  | 52        |
| 3.3.2    | Die 3'-UTR der MtH <sub>550</sub> -cDNA.....                                  | 53        |
| 3.3.3    | Die MtH <sub>550</sub> -cDNA besitzt ein Signalpeptid .....                   | 53        |
| 3.3.4    | Die MtH <sub>550</sub> -cDNA besitzt eine „Translation initiation site“ ..... | 54        |
| 3.3.5    | Die MtH <sub>550</sub> -Untereinheit besitzt sieben Varianten der FU-f.....   | 55        |
| 3.4      | Vergleich der FUs von MtH <sub>400</sub> und MtH <sub>550</sub> .....         | 57        |
| 3.4.1    | FU-a .....  | 57        |
| 3.4.2    | FU-b .....  | 58        |
| 3.4.3    | FU-c .....  | 59        |
| 3.4.4    | FU-d.....   | 60        |
| 3.4.5    | FU-e .....  | 61        |
| 3.4.6    | FU-f.....   | 62        |
| 3.4.7    | FU-g.....   | 63        |
| 3.4.8    | FU-h.....   | 64        |
| 3.4.9    | FU-f <sub>1</sub> bis FU-f <sub>6</sub> .....                                 | 65        |
| 3.4.10   | Statistische Daten der Untereinheiten des MtH .....                           | 67        |
| 3.5      | Aminosäurezusammensetzung des MtH .....                                       | 69        |

|          |   |            |
|----------|---|------------|
| <b>4</b> | <b>Diskussion</b> .....   | <b>71</b>  |
| 4.1      | „Codon-Usage“ Analyse der MtH-Untereinheiten .....                        | 73         |
| 4.2      | Vergleichende Sequenzanalyse des MtH .....                                | 75         |
| 4.3      | Analyse der potentiellen N-Glykosylierungsstellen .....                   | 82         |
| 4.4      | Analyse der Dekamer-Dekamer Interfaces .....                              | 84         |
| 4.5      | Evolution des Mega-Dekamers .....   | 87         |
| 4.5.1    | Entstehung der sechs C-terminalen FU-f Varianten .....                    | 89         |
| 4.6      | Molekulare Uhr .....  | 94         |
| <b>5</b> | <b>Zusammenfassung</b> .....  | <b>98</b>  |
|          | <b>Anhang</b> .....   | <b>99</b>  |
| <b>A</b> | <b>Abkürzungsverzeichnis</b> .....  | <b>99</b>  |
| A.1      | Allgemeine Abkürzungen.....   | 99         |
| A.2      | Abkürzungen in der Hämocyaninforschung .....                              | 101        |
| <b>B</b> | <b>IUPAC-Nomenklatur für Aminosäuren</b> .....                            | <b>102</b> |
| <b>C</b> | <b>IUPAC-Nomenklatur für Nukleotide</b> .....                             | <b>102</b> |
| <b>D</b> | <b>Der genetische Code</b> .....  | <b>103</b> |
| <b>E</b> | <b>Programme</b> .....  | <b>104</b> |
| <b>F</b> | <b>Primer</b> .....   | <b>105</b> |
| F.1      | Vektorspezifische Primer .....  | 105        |
| F.2      | Oligo-(dT)-Primer .....   | 105        |
| F.3      | Degenerierte Hämocyanin-Primer .....                                      | 105        |
| F.4      | Genspezifische Hämocyanin-Primer (MtH <sub>400</sub> -Untereinheit) ..... | 106        |
| F.5      | Genspezifische Hämocyanin-Primer (MtH <sub>550</sub> -Untereinheit) ..... | 108        |
| <b>G</b> | <b>cDNA-Sequenzen</b> .....   | <b>111</b> |
| G.1      | MtH <sub>400</sub> -Untereinheit .....                                    | 111        |
| G.2      | MtH <sub>550</sub> -Untereinheit .....                                    | 116        |
| <b>H</b> | <b>Alignment der FUs verschiedener Mollusken-Hämocyanine</b> .....        | <b>123</b> |
| <b>I</b> | <b>Prozentuale Aminosäurezusammensetzung des MtH</b> .....                | <b>128</b> |
|          | <b>Literaturverzeichnis</b> .....   | <b>130</b> |
|          | <b>Danksagung</b> .....   | <b>141</b> |
|          | <b>Curriculum Vitae</b> .....   | <b>142</b> |
|          | <b>Eidesstattliche Erklärung</b> .....                                    | <b>144</b> |

# 1 Einleitung

## 1.1 Höheres Leben benötigt Sauerstoff

Das erste Leben auf der Erde entwickelte sich in Form von obligat anaeroben Mikroorganismen. Mit der Entstehung der Photosynthese durch Cyanobakterien kam es schrittweise zu einer vermehrten Produktion von freiem Sauerstoff, welcher jedoch durch organisches Material und freie, in reduzierter Form vorliegende Metalle in den Ozeanen abgefangen wurde (*Review: Decker & van Holde, 2011*). Erst nach der vollständigen Oxidation dieser Metalle konnte sich die Konzentration des Sauerstoffs in der frühen Atmosphäre der Erde langsam erhöhen. Dieses als „Great Oxidation Event“ bezeichnete Ereignis liegt etwa 2,4 Milliarden Jahre zurück (Kump, 2008; *Review: Sessions, et al., 2009*). In der Folge entwickelten sich die ersten prokaryotischen Aerobier sowie die ersten aeroben Eukaryoten, die durch ihre effektivere oxidative Energiegewinnung einen selektiven Vorteil gegenüber den anaeroben Lebewesen aufweisen konnten. Mit der Evolution der frühen mehrzelligen Organismen vor etwa 1,7 bis 1 Milliarden Jahren begann sich das Leben auf der Erde weiter zu entwickeln und auszubreiten. Vor etwa 550 Millionen Jahren schließlich ereignete sich die äußerst kontrovers diskutierte „Kambrische Explosion“, im Rahmen derer viele der heute bekannten großen Tierstämme innerhalb eines kurzen geologischen Zeitraumes in den fossilen Schichten erstmals auftraten (Morris, 2006). Die frühen aeroben Metazoa deckten ihren Sauerstoffbedarf ausschließlich über Diffusion und waren daher auf einen sehr kleinen Körper bzw. eine dünne (1-2 mm) und längliche Körperform begrenzt (Raff *et al.*, 1970). Erst durch die Entwicklung von spezialisierten Gasaustauschflächen bzw. -organen, Kreislaufsystemen und freien oder zellgebundenen, Sauerstoff transportierenden Proteinen wurden die Größen- und Formbegrenzungen der Metazoa aufgehoben und so deutlich komplexere Lebensformen ermöglicht.

## 1.2 Respiratorische Proteine

Durch die Entstehung respiratorischer Proteine war es möglich, die Sauerstofftransportkapazität des Blutes bzw. der Hämolymphe im Vergleich zu rein physikalisch gelöstem Sauerstoff um ein Vielfaches zu erhöhen. Diese Proteine liegen

entweder in Form von großen Partikeln frei gelöst in der Hämolymphe oder als kleinere Partikel in Zellen verpackt vor.

Heute sind drei Klassen Sauerstoff transportierender Proteine bekannt. Man unterscheidet die Hämerythrine, die Hämoglobine und die Hämocyanine. Die Hämerythrine sind als respiratorische Proteine auf wenige Invertebraten beschränkt und wurden bisher nur bei Sipunculiden, Brachiopoden, Priapuliden und Anneliden beobachtet. Sie verfügen über ein binukleäres Eisen(II)-Zentrum und sind auf ein intrazelluläres Vorkommen beschränkt (*Review*: Kurtz, 1992; Bailly *et al.*, 2008; Meyer *et al.*, 2010).

Hämoglobine hingegen sind ein weitverbreiteter Typ von intra- und extrazellulär auftretenden respiratorischen Proteinen. Ihr Vorkommen erstreckt sich nahezu über das gesamte Tierreich. Das aktive Zentrum der Hämoglobine verfügt über ein Porphyrin-ringsystem, welches über mehrere Histidinreste mit dem Proteinanteil verknüpft ist. Im Zentrum des Porphyrinrings befindet sich das über die Stickstoffreste der Pyrrolringe komplexierte Eisen(II)-Atom, welches den Sauerstoff reversibel bindet (*Reviews*: Hardison, 1996; Terwilliger, 1998; Weber & Vinogradow, 2001).

Die Hämocyanine, bei welchen es sich um die vorherrschenden Sauerstoff transportierenden Proteine der Arthropoden und Mollusken handelt, bilden die dritte Klasse der respiratorischen Proteine (*Review*: van Holde & Miller, 1995). Das aktive Zentrum der Hämocyanine beinhaltet zwei Sauerstoff bindende Kupfer(I)-Atome, welche jeweils an drei hochkonservierte Histidin-Reste gebunden vorliegen (Abb. 1a). Durch die Bindung an Sauerstoff (in Form von  $O_2^{2-}$ ) werden die beiden Cu(I) Atome formal zu Cu(II) oxidiert. Diese reversible Bindung von Sauerstoff bewirkt bei dem ansonsten farblosen Protein eine tiefblaue Färbung, welche auch Ursprung des Namens „Hämocyanin“ ist (*Review*: van Holde & Miller, 1995; Markl, 1996; Cuff *et al.*, 1998).

Arthropoden-Hämocyanine unterscheiden sich strukturell deutlich von den Mollusken-Hämocyaninen (*Reviews*: Markl & Decker, 1992; van Holde & Miller, 1995). Es handelt sich um globuläre Proteine, die als Hexamere oder Oligohexamere vorliegen und aus 75 kDa Untereinheiten aufgebaut sind. Jede der aus drei strukturellen Domänen bestehenden Untereinheiten besitzt ein aktives Kupferzentrum und kann somit ein Sauerstoffmolekül reversibel binden. Details zur Struktur der Arthropoden-Hämocyanine finden sich z.B. bei (Burmester, 2001; 2002; Martin *et al.*, 2007).

## 1.2.1 Mollusken-Hämocyanine: Eine Übersicht

Die Grundeinheit des Mollusken-Hämocyanins ist die globuläre funktionelle Einheit (functional unit, FU). Sie weist zwei distinkte strukturelle Domänen auf (Cuff *et al.*, 1998). Die größere N-terminale „Core-Domäne“ besteht vorwiegend aus  $\alpha$ -Helices und beinhaltet das binukleäre aktive Zentrum mit den beiden als Cu-A und Cu-B bezeichneten Kupferatomen (Abb. 1a und b). Die „ $\beta$ -Sandwich Domäne“ am C-terminalen Ende besteht hauptsächlich aus  $\beta$ -Faltblatt Motiven (Abb. 1b) (Cuff *et al.*, 1998; Perbandt *et al.*, 2003; Jaenicke *et al.*, 2011).

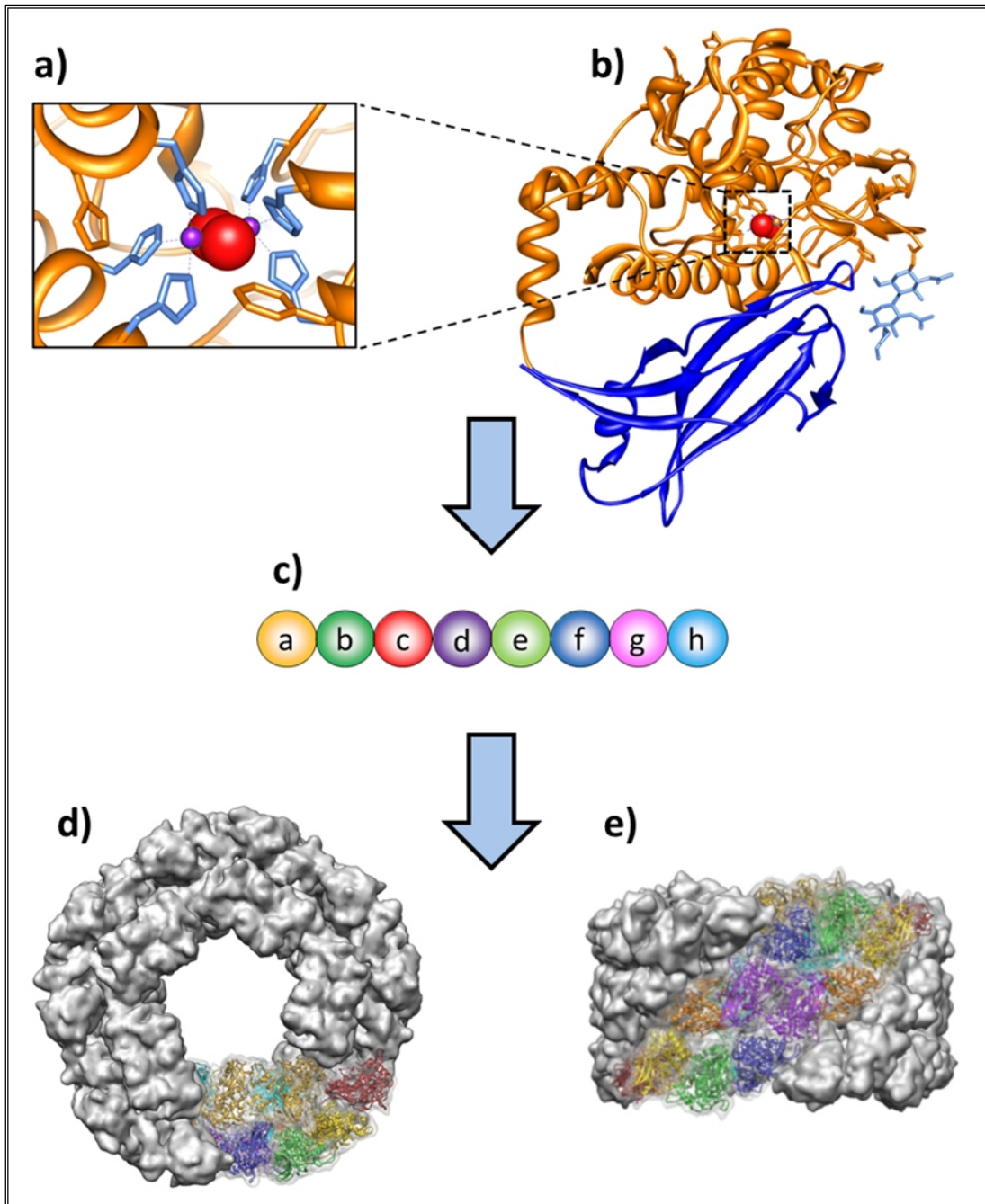
Das aktive Zentrum mit seinen beiden durch sechs Histidine gebundenen Kupferatomen gehört zur Gruppe der Typ-III-Kupferproteine und ähnelt den ebenfalls in dieser Gruppe beheimateten Tyrosinasen und Catechol-Oxidasen (van Holde & Miller, 1995; Klabunde *et al.*, 1998). Man geht heute davon aus, dass sich Hämocyanine, Catechol-Oxidasen und Tyrosinasen unabhängig voneinander aus einem primitiven Vorläufer-Typ-III-Kupferprotein entwickelt haben (*Reviews*: van Holde *et al.*, 2001; Jaenicke & Decker, 2004; Decker *et al.*, 2007).

Sieben bis acht konkatenierte, funktionelle Einheiten (FU) bilden die 350 bis 450 kDa großen Untereinheiten des Mollusken-Hämocyanins (Abb. 1c). Die jeweils 45 bis 55 kDa großen FUs sind untereinander durch 10 bis 20 Aminosäuren lange „Linker“-Peptide verbunden und werden vom N-terminalen Ende beginnend als FU-a bis FU-h bezeichnet (Lang *et al.*, 1991; Lieb *et al.*, 2001).

Das typische halbhohle zylinderförmige Hämocyanin-Dekamer mit 35 nm Durchmesser entsteht durch die Zusammenlagerung von zehn Kopien der Untereinheit und besitzt eine Molekularmasse von etwa 4 MDa (Abb. 1d und e). Die Untereinheiten lagern sich innerhalb des Dekamers paarweise als asymmetrische Homodimere zusammen (Gatsogiannis & Markl, 2009). Die Zylinderwand wird durch die FU-a bis FU-f gebildet, während die FU-g und FU-h den einseitigen Kragenkomplex („collar-arc-Komplex“) bilden (Siezen & van Bruggen, 1974; Lang, 1988; Lang & van Holde, 1991; van Holde & Miller, 1995; Lieb *et al.*, 1999; Meissner *et al.*, 2000; 2007; Gatsogiannis & Markl, 2009).

Abweichend von der oben beschriebenen Struktur der funktionellen Einheiten des Mollusken-Hämocyanins verfügt die FU-h über einen zusätzlichen 100 Aminosäuren umfassenden C-terminal gelegenen „Tail“ (Lieb *et al.*, 2000). Dieser Schwanzteil weist eine hohe Ähnlichkeit zur Struktur des „cucumber basic protein“ (CBP) auf, welches zur Gruppe der Typ-I-Kupferproteine gehört (Jaenicke *et al.*, 2010; 2011). Der zusätzliche kupferlose „Tail“ der FU-h besitzt ein Faltungsmotiv ähnlich der Cupredoxine und übt

wahrscheinlich die Funktion eines „Kupfer-Chaperons“ aus, welches ursprünglich für die Kupferbeladung des aktiven Zentrums des Vorläuferproteins verantwortlich war (Jaenicke, *et al.*, 2010; 2011).

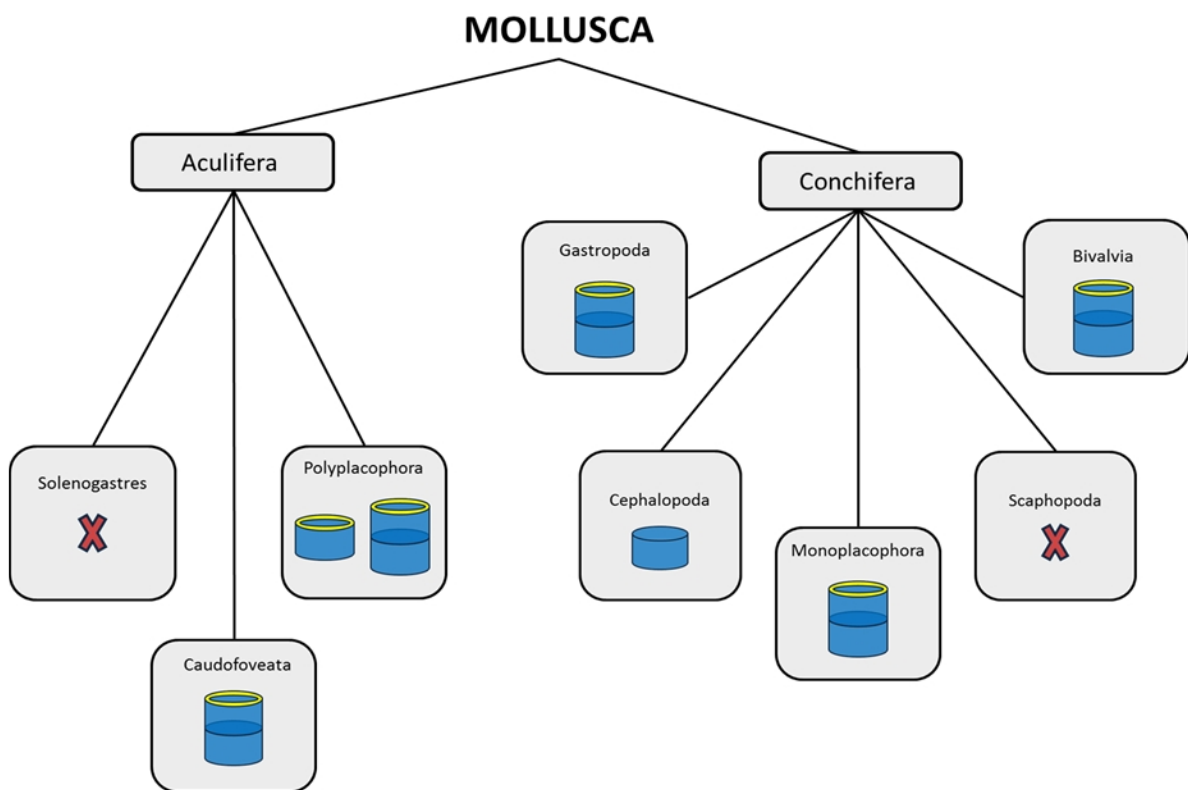


**Abb. 1: Aufbau eines Mollusken-Hämocyanins am Beispiel des KLH1**

(a) Aktives Zentrum eines Mollusken-Hämocyanins mit den beiden Kupferatomen (violett) sowie dem gebundenen Sauerstoffmolekül (rot). Die sechs kupferbindenden Histidine (cyan) sind ebenfalls erkennbar. (b) Röntgenstruktur einer funktionellen Einheit des Hämocyanins. Die „Core“-Domäne (orange) mit dem aktiven Zentrum (schwarzer Rahmen) sowie die „β-Sandwich“-Domäne (blau) sind deutlich erkennbar (PDB-ID-Code: 1JS8; verändert nach Cuff *et al.*, 1998). (c) Schema der aus acht FUs bestehenden Untereinheit des Hämocyanins. Jeweils zwei Untereinheiten lagern sich zu asymmetrischen Homodimeren zusammen. Top- (d) und Seitenansicht (e) des aus fünf Homodimeren bestehenden zylinderförmigen Dekamers (3D-Modell verändert nach Lieb *et al.*, 2010).

Bisher konnte das Hämocyanin in sechs der acht rezenten Klassen der Mollusca nachgewiesen werden (Abb. 2).

Im Falle der Gastropoden und einiger Bivalvia treten die Hämocyanine als Didekamere oder Multidekamere auf. Dabei assemblieren zwei asymmetrische Dekamere zu einem Didekamer, indem sie sich derart aneinander lagern, dass die Kragenregionen jeweils nach außen zeigen. Multidekamere bilden sich, indem sich weitere Dekamere an die Enden eines zentralen Didekamers anlagern (Herskovitz & Hamilton, 1991; Markl *et al.*, 1991; van Holde & Miller, 1995; Markl *et al.*, 2001; Lieb *et al.*, 2004; 2008; 2010; Bergmann *et al.*, 2007).



**Abb. 2: Vorkommen der Hämocyanine im Stamm Mollusca**

Die Abbildung zeigt symbolisch das Vorkommen der Hämocyanine in den acht rezenten Klassen (nach Steitz & Stängel, 1984) des Stammes Mollusca. Ein blauer Zylinder mit gelbem Kragen symbolisiert das aus den funktionellen Einheiten FU-a bis FU-h aufgebaute Dekamer. Das aus zwei Dekameren bestehende Didekamer wird durch zwei gestapelte blaue Zylinder dargestellt. Der Zylinder ohne Kragen zeigt das aus den FU-a bis FU-g bestehende Dekamer der Cephalopoden. Die roten Kreuze bezeichnen Klassen, in denen bisher kein Hämocyanin nachgewiesen werden konnte.

Die Cephalopoden bilden ausschließlich Dekamere, deren Untereinheiten aus sieben (*Octopus*-Typ: FU-a bis FU-g) oder acht (*Sepia*-Typ: FU-a bis FU-g + FU-d') funktionellen Einheiten bestehen. Durch das Fehlen der FU-h besitzen sie keinen äußeren

sondern nur den durch die FU-g gebildeten inneren Kragen (Herskovitz, 1988; Cuff *et al.*, 1998; Miller *et al.*, 1998; Bergmann *et al.*, 2006; Gatsogiannis *et al.*, 2007).

Das Hämocyanin der Monoplacophoren besteht aus den funktionellen Einheiten FU-a bis FU-h und ist vermutlich in der Lage Didekamere auszubilden (Lieb *et al.*, unveröffentlichte Daten).

Bei Scaphopoden und Solenogastren konnten Hämocyanine bisher nicht nachgewiesen werden. Stattdessen dient bei beiden Klassen das Hämoglobin als respiratorisches Protein (Lieb *et al.*, 2008).

Die Klasse der Caudofoveata hingegen verfügt über Hämocyanin. Durch molekulargenetische Untersuchungen konnten cDNA-Sequenzen der FU-e, FU-f, FU-g und FU-h amplifiziert werden (Lieb *et al.*, 2008). Es ist daher wahrscheinlich, dass die Hämocyanin-Untereinheiten der Caudofoveata die typischen funktionellen Einheiten FU-a bis FU-h umfassen und sowohl Dekamere als auch Didekamere bilden können.

Die Polyplacophora schließlich besitzen entweder ausschließlich solitäre Dekamere (Chitonida) oder Dekamere und Didekamere (Lepidopleurida) (Möller, 2008).

## 1.2.2 Das Hämocyanin der Gastropoden

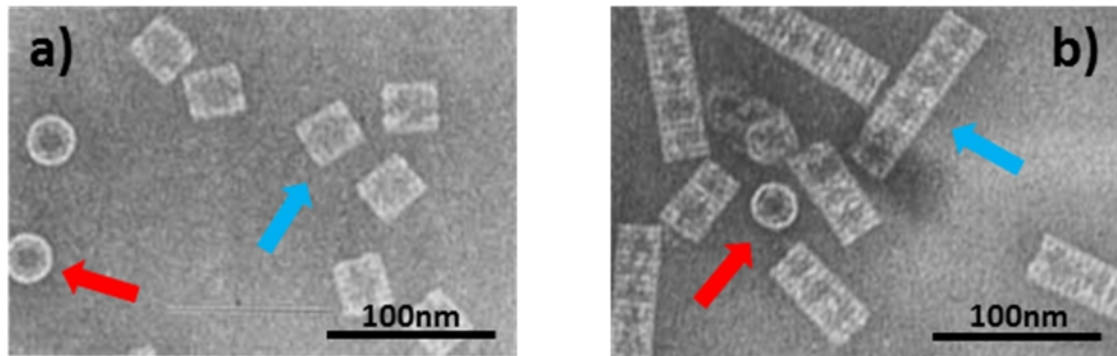
Die Untereinheit des Gastropoden-Hämocyanins besteht aus den acht funktionellen Einheiten FU-a bis FU-h und lagert sich vorwiegend zu Didekameren mit einer Molekularmasse von etwa 8 MDa zusammen (Abb. 3a). Dabei assemblieren zwei C<sub>5</sub>-symmetrische Dekamere mit den offenen Seiten, während die kragentragenden Seiten nach außen zeigen. Außer Dekameren und Didekameren treten bei Gastropoden auch tubuläre Multidekamere auf (Abb. 3b). (Herskovitz & Hamilton, 1991; Markl *et al.*, 1991; van Holde & Miller, 1995; Gatsogiannis & Markl, 2009; Lieb *et al.*, 2010).

An der Bildung der Didekamere sind FU-a, FU-b, FU-c und FU-f beider Dekamere beteiligt. Zwischen diesen FUs konnten mit Hilfe eines auf einer 3D-Rekonstruktion basierenden molekularen Modells vier unterschiedliche Interface-Typen bzw. Kontaktstellen identifiziert werden (Gatsogiannis & Markl, 2009). Zusätzlich wurden im Rahmen dieser Studie 23 weitere Interfaces zwischen den funktionellen Einheiten des Hämocyanins von *Megathura crenulata* (Keyhole limpet hemocyanin; KLH1) beschrieben. Man geht davon aus, dass diese Interfaces sowohl der Gesamtstabilität dienen, als auch an allosterischen Wechselwirkungen der FUs untereinander beteiligt sind (Gatsogiannis & Markl, 2009).

Viele der bisher untersuchten Gastropoden-Hämocyanine weisen eine deutliche Heterogenität ihrer Untereinheiten auf. Sowohl *Haliotis tuberculata* (Keller *et al.*, 1999; Lieb *et al.*, 1999; 2000) als auch *Megathura crenulata* (Senozan *et al.*, 1981; Markl *et al.*, 1991; Gebauer *et al.*, 1994; Söhngen *et al.*, 1997) und *Rapana thomasiana* (Idakieva *et al.*, 1993; Gebauer *et al.*, 1999) besitzen mindestens zwei verschiedene und zumindest partiell immunologisch distinkte Hämocyanin Untereinheiten. Im Falle des KLH unterscheiden sich die beiden Isoformen KLH1 und KLH2 sowohl biochemisch als auch immunologisch und werden differentiell exprimiert. Ein auffallendes Beispiel dafür sind die unterschiedlichen Quartärstrukturen der beiden Untereinheiten. KLH1 bildet Cluster aus Didekameran, während KLH2 die erwähnten tubulären Multidekamere ausbildet (Harris *et al.*, 2000). Das Verhältnis der beiden Isoformen variiert von Individuum zu Individuum. Außerdem wurde beobachtet, dass bei Tieren in Gefangenschaft nach einigen Monaten ausschließlich KLH2 vorhanden war und KLH1 anscheinend vollkommen abgebaut wurde (Markl *et al.*, 1991; Gebauer *et al.*, 1994). Auch im Falle des *Haliotis tuberculata* Hämocyanins (HtH) wurde ein individuell unterschiedliches Verhältnis der Isoformen HtH1 und HtH2 beobachtet. Bei den beiden Untereinheiten der Gastropoden-Hämocyanine handelt es sich also um unterschiedlich regulierte Proteine mit deutlich verschiedenen Halbwertszeiten in der Hämolymphe und unterschiedlichem Assemblierungsverhalten. An den beschriebenen Unterschieden zwischen den Untereinheiten könnte auch das deutlich abweichende Glykosylierungsmuster der beiden Untereinheiten verantwortlich sein (Keller *et al.*, 1999; Lieb *et al.*, 1999; 2000). Bei den bisherigen Untersuchungen der Gastropoden-Hämocyanine konnten verschiedene potentielle N-Glykosylierungsstellen identifiziert werden. Im Falle des KLH1 wurden acht Sequenzmotive gefunden, die dem Muster NXT/S entsprechen (Gatsogiannis & Markl, 2009). Das Hämocyanin von *Aplysia californica* besitzt sieben potentielle N-Glykosylierungsstellen (Lieb *et al.*, 2004). Bei Untersuchungen des HtH1 wurden sogar 13 dieser Motive identifiziert (Lieb *et al.*, 2000). Den drei aufgeführten Gastropoden-Hämocyaninen gemein ist das Fehlen der N-Glykosylierungsstellen innerhalb der FU-c. Ob und welche unterschiedlichen physiologischen Funktionen die beiden Isoformen innehaben, konnte bisher nicht geklärt werden.

Der Ort der Biosynthese des Hämocyanins konnte für den marinen Gastropoden *Haliotis tuberculata* durch immunhistochemische Untersuchungen in Verbindung mit „*in-situ*“-Hybridisierungsexperimenten abschließend geklärt werden (Streit, 2008). Die Rhogozyten oder Poren-Zellen des Bindegewebes verschiedener Körperbereiche konnten so als Ort der

Hämocyaninsynthese identifiziert werden (Albrecht *et al.*, 2001). Bei zeitgleich an *Megathura crenulata* durchgeführten Untersuchungen wurden zwar auch Rhogozyten identifiziert, diese enthielten jedoch kein Hämocyanin. Diese Beobachtung wird mit einer saisonalen Expression des Hämocyanins in diesen Tieren in Verbindung gebracht (Albrecht *et al.*, 2001).



**Abb. 3: Transmissions-Elektronenmikroskopische Aufnahmen des KLH**

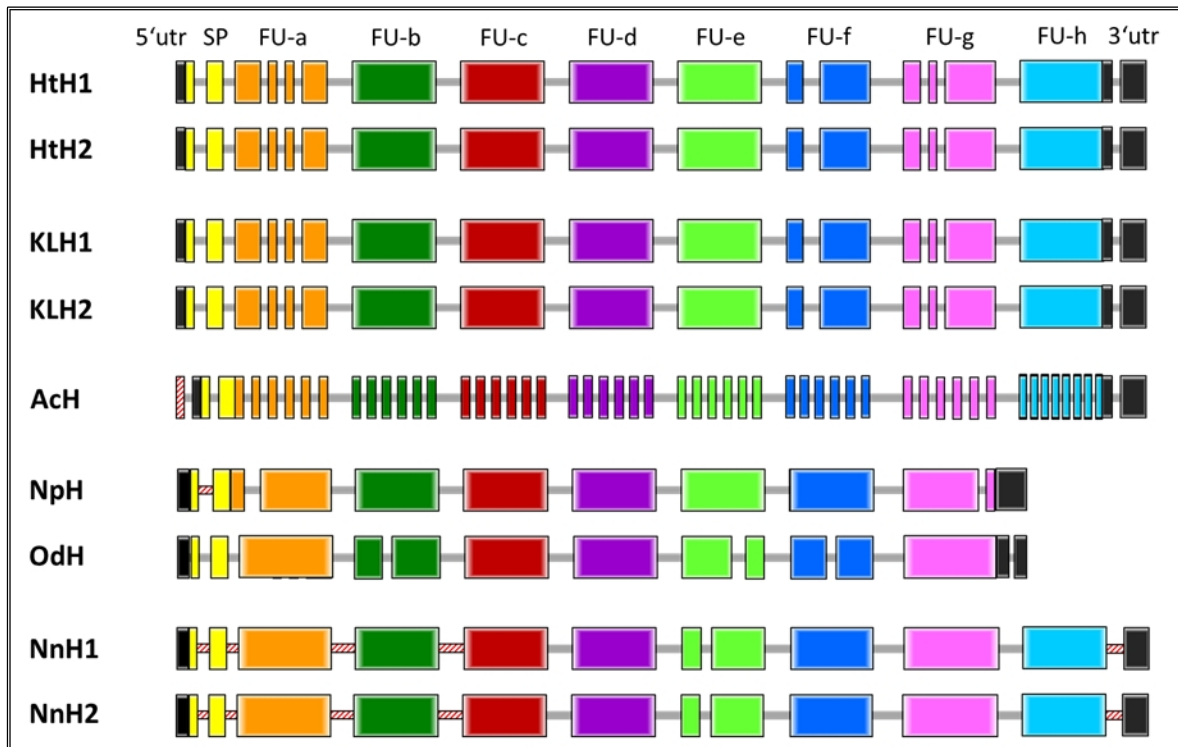
Aufgereinigte KLH1 Didekamere (**a**) sowie Multidekamere (**b**) jeweils in Top- (roter Pfeil) und Seitenansicht (blauer Pfeil) (verändert nach Söhngen *et al.*, 1997).

### 1.2.3 Die Genstruktur der Mollusken-Hämocyanine

Im letzten Jahrzehnt wurden die Hämocyanine verschiedenster Mollusken auf cDNA sowie auf genomischer DNA-Ebene sequenziert. Vergleiche der bisher bekannten Sequenzen zeigen eine hohe Ähnlichkeit der Genstruktur innerhalb des gesamten Stammes der Mollusken (siehe Abb. 4).

Der in Abb. 4 dargestellte Vergleich verschiedener Hämocyanin-Gene zeigt ihre strukturellen Übereinstimmungen innerhalb der Mollusken. Deutlich erkennbar ist die charakteristische Unterteilung in die sieben bis acht verschiedenen funktionellen Einheiten sowie die dazwischen liegenden über die Klassen hinweg konservierten Linker-Introns.

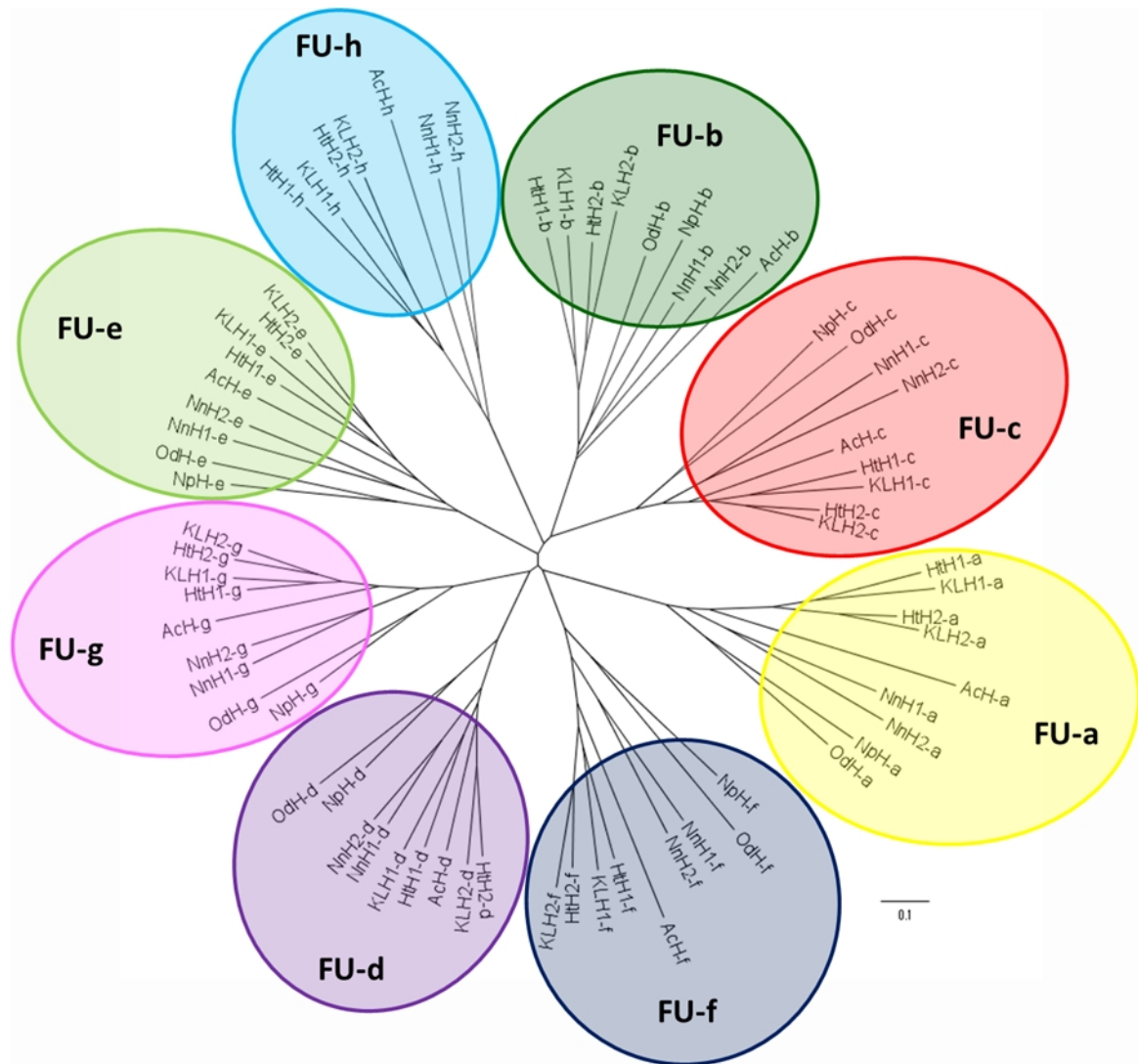
Die kodierenden Abschnitte der FU-a bis FU-g bzw. FU-h bilden eine lineare Sequenz und werden durch die Phase 1 Linker-Introns unterschiedlicher Länge getrennt (Lieb *et al.*, 2001). Zusätzlich existieren in den Genen sogenannte interne Introns, die innerhalb der Exonbereiche inserieren. Anzahl und Lage der internen Introns variieren deutlich zwischen den bisher untersuchten Spezies und sind nicht innerhalb des Stammes Mollusca konserviert. Ein dritter, sehr variabler Intronotyp inseriert in den beiden untranslatierten Regionen des Gens sowie im Bereich des Signalpeptids (Lieb *et al.*, 2001; 2004). Ein Signalpeptid ist bei allen untersuchten Mollusken-Hämocyanin Genen vorhanden.



**Abb. 4: Strukturvergleich verschiedener Mollusken-Hämocyanin Gene**

Die Grafik zeigt vergleichend die Hämocyaningene der Gastropoden *Haliothis tuberculata* (HtH1, HtH2), *Megathura crenulata* (KLH1, KLH2) und *Aplysia californica* (AcH) der Cephalopoden *Nautilus pompilius* (NpH) und *Octopus dofleini* (OdH) sowie der protobranchen Muschel *Nucula nucleus* (NnH1 und NnH2). Die farbliche Kennzeichnung der einzelnen Strukturen ist wie folgt: utr = schwarz, Signalpeptid = gelb, FU-a = orange, FU-b = dunkelgrün, FU-c = rot, FU-d = violett, FU-e = hellgrün, FU-f = dunkelgrün, FU-g = rosa, FU-h = hellblau. Die Linker-Introns sowie die internen Introns sind durch graue Balken gekennzeichnet. Rot schraffierte Regionen sind bisher nicht sequenziert. Es ist deutlich erkennbar, dass sich alle dargestellten Gene in ihrem grundsätzlichen Aufbau sehr ähnlich sind. Die beiden Isoformen von HtH, KLH und NnH stimmen jeweils strukturell exakt überein. Deutlich erkennbar ist auch die fehlende FU-h bei den beiden dargestellten Cephalopoden-Hämocyanin Genen (verändert nach Altenhein, 2000; Lieb *et al.*, 2001; 2004; Bergmann *et al.*, 2006; 2007).

Bei phylogenetischen Vergleichen der kodierenden Hämocyaninsequenzen verschiedener Mollusken-Spezies wurde beobachtet, dass die orthologen FUs immer zusammen in einem Ast des entsprechenden Stammbaumes gruppieren (Lieb *et al.*, 2001; 2004). Ein entsprechender Stammbaum, der auf dem Alignment der funktionellen Einheiten von neun verschiedenen Hämocyanin-Genen basiert, ist in Abb. 5 dargestellt. Aufgrund dieser Beobachtungen geht man davon aus, dass sich die heutigen Hämocyanin-Gene aus einer einzigen Vorläufer-FU durch drei aufeinanderfolgende Genduplikationen und Fusionen entwickelt haben. Dieses Ereignis muss dementsprechend vor der Radiation der Mollusken in die rezenten Klassen stattgefunden haben (Lieb *et al.*, 2001; Lieb & Markl, 2004).

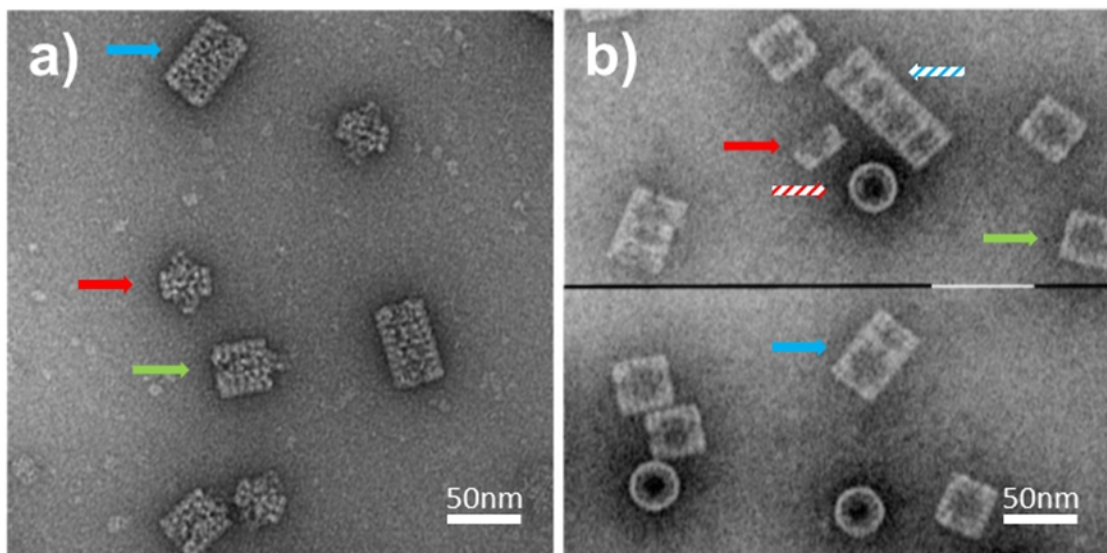


**Abb. 5: Radialer Stammbaum der funktionellen Einheiten verschiedener Mollusken-Hämocyanine**

Die orthologen FUs der verschiedenen Hämocyanine gruppieren innerhalb des radialen Stammbaums jeweils zusammen in einem der acht Äste. Für die Berechnungen wurden die abgeleiteten Aminosäuresequenzen der einzelnen FUs der Hämocyanine von *Haliotis tuberculata* (HtH1 und HtH2; Lieb *et al.*, 2000; 2001a; Altenhein *et al.*, 2002), *Megathura crenulata* (KLH1 und KLH2; Streit *et al.*, 2003), *Aplysia californica* (AcH; Boisguérin *et al.*, 2003), *Nautilus pompilius* (NpH; Bergmann *et al.*, 2006), *Octopus dofleini* (OdH; Miller *et al.*, 1998; Lieb *et al.*, 2001a) und *Nucula nucleus* (NnH1 und NnH2; Bergmann *et al.*, 2007) mit Hilfe des Programms „MUSCLE“ aligniert. Aus dem multiplen Sequenzalignment wurde mit dem Neighbour-Joining-Algorithmus ein phylogenetischer Stammbaum berechnet. Als Substitutionsmodell wurde das „Jones-Taylor-Thornton Modell“ (JTT) verwendet (FU-a = gelb, FU-b = dunkelgrün, FU-c = rot, FU-d = violett, FU-e = hellgrün, FU-f = dunkelblau, FU-g = rosa, FU-h = hellblau).

### 1.3 Das Mega-Hämocyanin der Cerithioidea

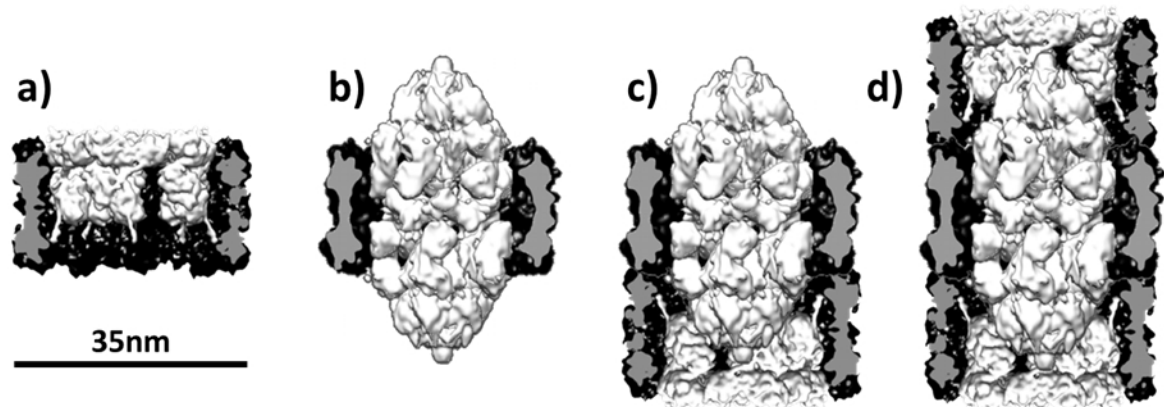
Während einer routinemäßigen Untersuchung von Schnecken der Art *Leptoxis carinata* wurden bei elektronenmikroskopischen Aufnahmen der Hämolymphe ausschließlich Hämocyanin-Tridekamere identifiziert (Lieb *et al.*, 2010). Dieses sehr ungewöhnliche Szenario war noch bei keinem der bisher untersuchten Hämocyanine beobachtet worden. Auch das Aussehen der Hämocyanin-Tridekamere unterschied sich deutlich von dem eines typischen Gastropoden-Hämocyanins. Der innere Bereich der Tridekamere war im Gegensatz zu den tubulären, teils hohlen Tridekameran anderer Gastropoden-Hämocyanine nahezu vollständig durch Masse ausgefüllt. In Abb. 6a ist eine elektronenmikroskopische Aufnahme (TEM) des Hämocyanins der zeitgleich untersuchten Schnecke *Melanooides tuberculata* (Cerithioidea) dargestellt. Deutlich erkennbar sind die Tridekamere sowie Didekamere, die in der Form von den regulären Didekameran (Abb. 6b) der Gastropoden-Hämocyanine abweichen. Die beobachteten Didekamere weisen an einem Ende eine deutlich hervortretende Masse auf (siehe auch Abb. 7c). Weiterhin konnten Dekamere mit zusätzlichen Massen an beiden Enden identifiziert werden (Abb. 6a; Abb. 7b) (Lieb *et al.*, 2010).



**Abb. 6:** Transmissions-Elektronenmikroskopische Aufnahmen des Hämocyanins von *Melanooides tuberculata* (a) sowie *Megathura crenulata* (b)

(a) Hämocyanin von *Melanooides tuberculata* bestehend aus Mega-Tridekameran (blauer Pfeil), Didekameran, die sich aus einem regulären Dekamer und einem Mega-Dekamer zusammensetzen (grüner Pfeil) sowie Mega-Dekameren (roter Pfeil). Einzelne Dekamere sind in dieser Probe nicht sichtbar, waren aber bei anderen Untersuchungen durchaus vorhanden. (b) Hämocyanin KLH2 von *Megathura crenulata*. Deutlich erkennbar sind Dekamere in der Seiten- (roter Pfeil) und Topansicht (rotgestreifter Pfeil). Außerdem sind Didekamere (grüner Pfeil) und tubuläre Multidekamere (blau bzw. blaugestreifter Pfeil) vorhanden (nach Lieb *et al.*, 2010).

Durch die in unserer Arbeitsgruppe durchgeführten weitergehenden elektronenmikroskopischen Untersuchungen konnten zwei distinkte Bestandteile der Tridekamere unterschieden werden. Das fortan als Mega-Hämocyanin bezeichnete Tridekamer besteht aus einem zentralen Mega-Dekamer, welches die beobachteten beidseitigen Erweiterungen aufweist, sowie zwei flankierenden typisch geformten Dekameren. Die erwähnten Dekamere sind  $C_5$ -symmetrisch und weisen den bei Gastropoden-Hämocyaninen typischen einseitigen Kragenkomplex auf (Lieb *et al.*, 2010). Proteinbiochemische Untersuchungen zeigten zwei immunologisch verschiedene Untereinheiten, aus denen das Dekamer sowie das Mega-Dekamer aufgebaut sind. Anhand von Reassemblierungsversuchen wurde eine 400 kDa große Polypeptidkette als Untereinheit des regulären Dekamers bestätigt (Lieb *et al.*, 2010).



**Abb. 7:** 3D-Modelle des Mega-Hämocyanins von *Leptoxis carinata*

Die Abbildung zeigt verschiedene Assemblierungsstufen des Mega-Hämocyanins als medianen Längsschnitt durch die Zylinderwand. Das Dekamer (a) und das Mega-Dekamer (b) sind die Einheiten aus denen sich das Mega-Didekamer (c) mit der einseitigen Erweiterung (MD + D) sowie das Mega-Tridekamer (d) (MD + 2xD) zusammensetzen (nach Lieb *et al.*, 2010).

Eine weitere etwa 550 kDa große Untereinheit bildet das Mega-Dekamer (Lieb *et al.*, 2010). Bislang konnte unsere Arbeitsgruppe das Mega-Hämocyanin bei limnischen und marinen Cerithioidea aus 17 Familien nachweisen (unveröffentlichte Daten). Außerhalb des Monophylums Cerithioidea (Strong *et al.*, 2011) wurde bisher kein Mega-Hämocyanin gefunden. Daher scheint es sich bei dem Besitz des Mega-Hämocyanins um eine Synapomorphie der Cerithioidea zu handeln (Lieb *et al.*, 2010).

## 1.4 Die Cerithioidea

In Abb. 8 ist die phylogenetische Einordnung der Cerithioidea sowie der in dieser Arbeit untersuchten Gattung *Melanoidea* dargestellt. Die Cerithioidea bilden eine monophyletische Gruppe (Strong *et al.*, 2011). Sie untergliedert sich weiter in drei ebenfalls monophyletische Großgruppen, die jeweils mehrere Familien umfassen. Der in Abb. 9 dargestellte Stammbaum zeigt die Phylogenie der Cerithioidea basierend auf den Untersuchungen von Strong *et al.* (2011).

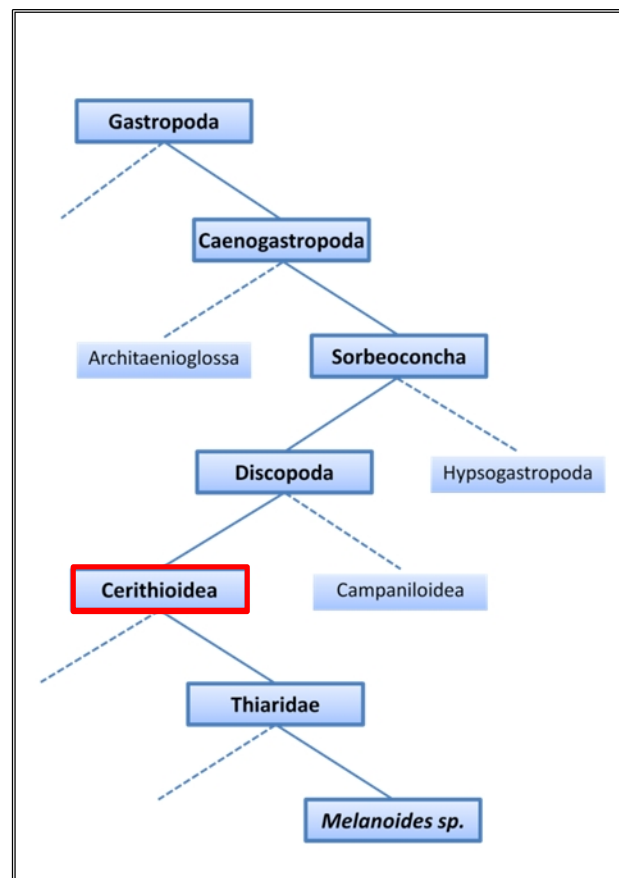
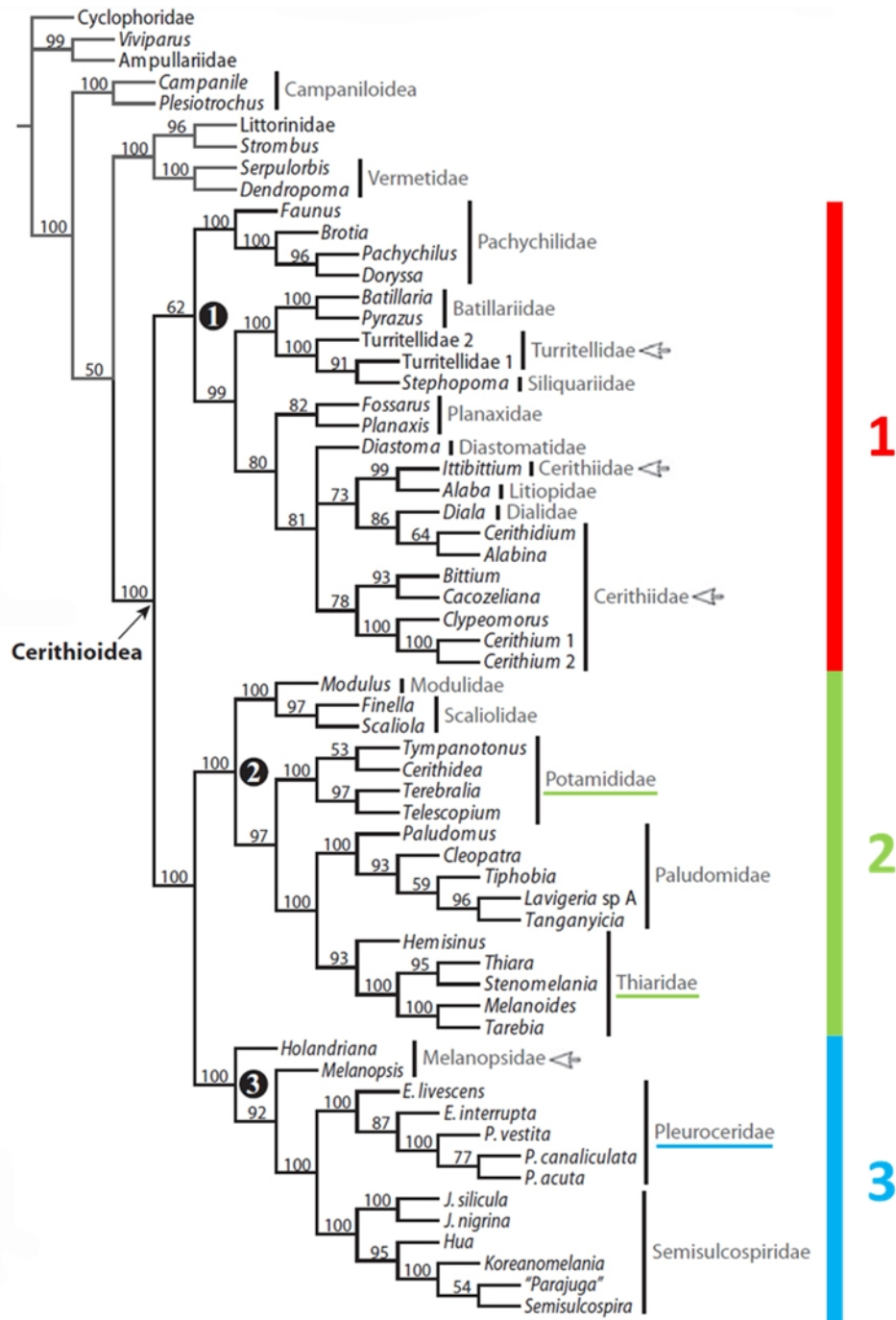


Abb. 8: Phylogenetische Einordnung der Überfamilie Cerithioidea sowie der in dieser Arbeit untersuchten Gattung *Melanoidea* (nach Tree Of Life web project; <http://tolweb.org>).

Innerhalb der Gastropoda stellen die Cerithioidea eine ebenso diverse wie artenreiche Gruppe dar (Starmühlner, 1969). Ihre Ausbreitung beschränkt sich meist auf subtropische und tropische Areale. Dort bevölkern sie eine Vielzahl an unterschiedlichsten Habitaten. Ihre Lebensräume umfassen Salzwasser-, Süßwasser- und Brackwasserregionen sowie Mangrovenwälder, Seegraswiesen, Flüsse, Bäche, Seen, Schlicktümpel und felsige Küstenbereiche (Houbrick, 1988). Manche Arten wie etwa *Melanoidea tuberculata* sind äußerst invasiv (Starmühlner, 1969).



**Abb. 9: Kombiniertes morphologisch-molekularer Stammbaum der Cerithioidea**

Der Stammbaum zeigt die Phylogenie der 18 rezenten Familien der Cerithioidea. Deutlich erkennbar ist der monophyletische Charakter der Überfamilie. Die drei ebenfalls monophyletischen Großgruppen innerhalb der Cerithioidea sind durch farbige Balken deutlich gekennzeichnet (verändert nach Strong *et. al.*, 2011). Gruppen, bei denen wir das Mega-Hämocyanin bereits genauer untersucht haben, sind unterstrichen (Lieb *et al.*, 2010).

## 1.5 Ziel der Arbeit

Nach der Entdeckung des Mega-Hämocyanins durch Lieb *et al.* (2010) war es das Ziel der vorliegenden Arbeit, die Primärstruktur beider Polypeptide aufzuklären. Die Sequenzdaten sollten im Folgenden statistisch ausgewertet und durch den Vergleich mit bereits bekannten Hämocyanindaten anderer Mollusken phylogenetisch analysiert werden.

Bei dem aus der 400 kDa Untereinheit gebildeten Dekamer handelte es sich anscheinend um eine typische Gastropoden-Hämocyanin-Untereinheit. Dies sollte anhand der Primärstruktur bestätigt werden. Weiterhin sollte geklärt werden, ob die Zylinderwand des Mega-Dekamers tatsächlich, wie von Lieb *et al.* (2010) vermutet, aus den typischen FU-a bis FU-f besteht. Von ganz besonderem Interesse war natürlich die Sequenz des Kragenbereiches des Mega-Dekamers. Gemessen an den 550 kDa der Untereinheit sollte dieser fünf bis sechs FUs umfassen.

Für die Versuche wurden Schnecken der Art *Melanoides tuberculata* verwendet, bei denen der Besitz des Mega-Hämocyanins bereits im Vorfeld durch elektronenmikroskopische und proteinbiochemische Untersuchungen nachgewiesen werden konnte (Lieb *et al.*, 2010). Die Tiere wurden schon vor Beginn der vorliegenden Arbeit in zwei unabhängigen Süßwasseraquarien des Instituts für Zoologie der Universität Mainz kultiviert und waren somit leicht verfügbar.

## 2 Material und Methoden

### 2.1 Tiere

Bei den im Rahmen dieser Arbeit untersuchten Tieren handelt es sich um Vertreter der Familie Thiariidae innerhalb der Überfamilie der Cerithioidea. Die Cerithioidea sind Teil der Caenogastropoda, welche eine Überordnung innerhalb der Klasse der Gastropoda darstellen.

#### 2.1.1 *Melanooides tuberculata*

*Melanooides tuberculata* (Malaiische Turmdeckelschnecke) ist eine zur Überfamilie der Cerithioidea gehörende Art, welche in Süßwasserhabitaten von Ostafrika bis Südostasien beheimatet ist. Durch ihre Beliebtheit bei Aquaristen ist sie jedoch heute auch vermehrt im europäischen Raum zu finden. Die allesfressenden Tiere sind vorwiegend nachtaktiv und meist im Bodensediment vergraben. *Melanooides tuberculata* ist getrenntgeschlechtlich und lebendgebärend. Die Fortpflanzung erfolgt dennoch in den meisten Fällen parthenogenetisch (Starmühlner, 1969; Houbrick, 1988; Stagl, 1993).



Abb. 10: *Melanooides tuberculata* (Bildquelle: Christoph Kühne)

Die zur Probengewinnung verwendeten Tiere der Art *Melanooides tuberculata* wurden in zwei unabhängig etablierten Süßwasseraquarien des Instituts für Zoologie der Universität Mainz bei einer Wassertemperatur von etwa 22°C artgerecht kultiviert. Für die entsprechenden Versuche wurden die Tiere jeweils frisch entnommen.

## 2.2 Chemikalien und Geräte

Sämtliche in dieser Arbeit verwendeten Chemikalien wurden in Analysequalität von den Firmen Martin Mundo OHG (Mainz), Merck (Darmstadt), Roche (Mannheim), Roth (Karlsruhe) und Sigma (Deisenhofen) bezogen. Für die Versuche wurden Petrischalen und Polypropylenröhrchen mit Schraubdeckel von Greiner (Heidelberg), 1,5 ml und 2 ml Reaktionsgefäße von Brand (Wertheim) sowie 0,2 ml PCR-Reaktionsgefäße von Thermo Fisher Scientific (Schwerte) verwendet. Wägearbeiten wurden mit Labor- oder Analysewaagen der Firma Mettler (Giessen) durchgeführt. Für sämtliche Zentrifugations-schritte wurden Tischzentrifugen von Eppendorf (Hamburg; Modell 5415R und 5424) sowie Starlab (Ahrensburg; Modell FugeONE) verwendet. Zentrifugationen von größeren Volumina wurden mit einer kühlbaren Megafuge 1.0R von Heraeus (Fulda) durchgeführt. Die verwendeten Wasserbäder und Heizblöcke stammten von den Firmen Julabo (Seelbach) sowie Techne<sup>®</sup> (Staffordshire, UK). Messungen des pH-Wertes wurden mit dem Digital-pH-Meter 646 der Firma Knick (Berlin) durchgeführt. Soweit nicht anders angeführt wurden photometrische Messungen mit dem Biophotometer der Firma Eppendorf (Hamburg) vorgenommen. Die verwendeten Brutschränke stammten von der Firma Memmert (Schwabach), während die Schüttler mit Inkubationshaube von der Firma Edmund Bühler GmbH (Hechingen) bezogen wurden. Gelelektrophoretische Auftrennungen wurden mit Elektrophoresekammern samt Zubehör der Firma Biotec Fischer (Reiskirchen) durchgeführt. Die verwendeten stabilisierten Netzteile des Typs E865 stammten von der Firma Consort (Turnhout, Belgien). Für die Auswertung und Dokumentation der Gele wurden UV-Transilluminatoren der Wellenlänge 312 nm (Vilber Lourmat, Eberhardzell) sowie 254 nm und 366 nm (Bachofer, Reutlingen) eingesetzt. Transport und Aufbewahrung von flüssigem Stickstoff wurden mit Dewargefäßen von KGW Isotherm (Karlsruhe) durchgeführt. Die Aufkonzentration von Nukleinsäurelösungen erfolgte mit dem SpeedVac Concentrator (Bachofer, Reutlingen). Auf darüber hinaus verwendete Materialien und Geräte wird in den entsprechenden Textpassagen hingewiesen.

## 2.3 Allgemeine Vorsichtsmaßnahmen

Sofern nicht abweichend beschrieben, wurden alle für molekular- und mikrobiologische Arbeiten verwendeten Lösungen, Geräte und Verbrauchsmaterialien bei 121°C und 1 bar

Überdruck für 30 Minuten autoklaviert. Die verwendeten Lösungen und Puffer wurden mit entmineralisiertem oder bidestilliertem Wasser angesetzt, welches den spezifischen Anforderungen entsprechend auch autoklaviert wurde. Bei allen durchgeführten Arbeitsschritten wurden Einmalhandschuhe getragen. Nukleinsäurelösungen sowie Enzyme wurden immer gekühlt bzw. tiefgekühlt gehalten und nur so lange wie nötig außerhalb von Gefrier- oder Kühlschränken aufbewahrt. Alle Arbeiten mit RNA wurden unter Verwendung von bidestilliertem, mit 0,1 % DEPC (Diethylpyrocarbonat) behandeltem Wasser, gestopften Pipettenspitzen, RNase freien Plastikwaren und frischen Handschuhen durchgeführt (siehe 2.5.1). Für Arbeiten, bei denen ein Kontakt mit Ethidiumbromid möglich war, wurden Nitrilhandschuhe getragen.

## 2.4 Mikrobiologische Methoden

### 2.4.1 Nährmedien und Agarplatten

|                                |   |
|--------------------------------|---|
| <b>LB-Medium</b> (pro Liter):  | 10 g NaCl<br>10 g Trypton<br>5 g Hefeextrakt<br>pH 7,0  |
| <b>LB-Agar</b> (pro Liter):    | 20 g Agar-Agar wurden vor dem Autoklavieren zu einem Liter LB-Medium gegeben.   |
| <b>SOC-Medium</b> (pro Liter): | 20 g Trypton<br>5 g Hefeextrakt<br>0,5 g NaCl<br>2,5 ml 1 M KCl<br>10 ml 1 M MgCl <sub>2</sub><br>20 ml 1 M Glukose<br>pH 7,0 |

Die aufgeführten Nährmedien wurden laut Beschreibung hergestellt und vor Verwendung bei 121°C und 1 bar Überdruck für 30 Minuten autoklaviert. Für die Herstellung von Agarplatten für selektives Bakterienwachstum wurde der autoklavierte LB-Agar nach dem Abkühlen auf etwa 50°C mit einem entsprechenden Antibiotikum versetzt:

**Ampicillin:** 100 mg / Liter LB-Medium

**Kanamycin:** 50 mg / Liter LB-Medium

Um eine Blau-Weiß-Selektion (siehe 2.5.9.1) zu ermöglichen, wurden dem LB-Agar vor dem Gießen der Platten noch folgende Mengen Isopropyl- $\beta$ -D-1-Thiogalactopyranosid (IPTG) sowie 5-Brom-4-Chlor-3-Indoyl- $\beta$ -D-Galactopyranosid (X-Gal) zugegeben:

**IPTG:** 0,5 mM

**X-Gal:** 80  $\mu$ g/ml

Der komplettierte LB-Agar wurde in 25 ml Portionen in sterile Petrischalen ( $\varnothing$  8,5 cm) gegossen. Nach dem vollständigen Erkalten wurden die Platten bei 4°C im Kühlraum gelagert. Die Lagerung der Nährmedien erfolgte lichtgeschützt bei Raumtemperatur.

## 2.4.2 Plattenkultur

Die Bakterien einer Transformationsreaktion oder einer Glycerinkultur wurden unter sterilen Bedingungen auf einer entsprechend antibiotikabehandelten Agarplatte ausgestrichen und über Nacht bei 37°C inkubiert. Anschließend wurde die Platte mit Parafilm<sup>®</sup> verschlossen und konnte so mehrere Monate bei 4°C gelagert werden.

## 2.4.3 Flüssigkultur

Zur Herstellung einer Flüssigkultur wurde eine Bakterienkolonie mithilfe einer sterilen Pipettenspitze von einer Plattenkultur aufgenommen und in ein 15 ml Kulturröhrchen überführt. Anschließend wurden 3 ml LB-Medium zugesetzt, welches zuvor mit einem, der Resistenz des Bakterienvektors entsprechenden, Antibiotikum versetzt worden war. Der so vorbereitete Kulturansatz wurde über Nacht bei 37°C und 200 rpm geschüttelt.

## 2.4.4 Glycerinkultur

Durch die Anlage von Glycerinkulturen war es möglich die Bakterien einer Flüssigkultur für einen längeren Zeitraum zu konservieren. Zur Herstellung einer Glycerinkultur wurden 500  $\mu$ l der Bakteriensuspension mit 500  $\mu$ l Glycerin vermischt, eingefroren und bei -80°C aufbewahrt.

## 2.4.5 Bakterienstämme und Vektoren

Im Rahmen dieser Arbeit wurden folgende Bakterienstämme und Vektoren verwendet:

### 2.4.5.1 Bakterienstämme:

- *Escherichia coli* StrataClone™ SoloPack® competent cells (Agilent, Böblingen)
- *Escherichia coli* One Shot® TOP10 chemically competent cells (Invitrogen, Darmstadt)

### 2.4.5.2 Vektoren

- StrataClone™ PCR Cloning Vector pSC-A-amp/kan (Agilent, Böblingen): 4,3 kb; Ampicillin- und Kanamycin-Resistenz; multiple cloning site (MCS); lacZ'  $\alpha$ -Gen (siehe Abb. 15)
- pCR-XL-TOPO® (Invitrogen, Darmstadt): 3,5 kb; Zeocin- und Kanamycin-Resistenz; multiple cloning site (MCS); ccdB lethal-Gen; LacZ $\alpha$ -Gen (siehe Abb. 14)

## 2.5 Molekularbiologische Methoden

### 2.5.1 Präparation von Nukleinsäuren

Zur Präparation der Nukleinsäuren wurden die Tiere jeweils kurz vorher aus dem Aquarium des Institutes entnommen und zur Betäubung etwa 30 Minuten auf Eis abgekühlt. Danach wurde die Schale vorsichtig entfernt und das Gewicht des Tieres bestimmt. Im Folgenden wurde das Tier in flüssigem Stickstoff (-196°C) schockgefroren und in einem autoklavierten Keramikmörser pulverisiert.

Für die Präparation von RNA galten besondere Vorsichtsmaßnahmen, um eine Kontamination mit ubiquitären RNasen zu vermeiden. Es wurden ausschließlich RNase-freie Materialien und Reagenzien, autoklavierte Gefäße und gestopfte Spitzen verwendet. Zudem wurde die Arbeitsfläche sowie alle verwendeten Geräte vor dem Gebrauch mit einer dekontaminierenden Lösung (RNase AWAY®; Invitrogen, Darmstadt) behandelt. Alle Arbeitsschritte wurden zügig und auf Eis durchgeführt, wobei ständig mit Handschuhen gearbeitet wurde. Für Arbeiten mit RNA wurde ausschließlich mit 0,1 % DEPC (Diethylpyrocarbonat) behandeltes und anschließend autoklaviertes Wasser

verwendet. Das DEPC wirkt dabei als unspezifischer RNase-Inhibitor, indem es mit Aminen, Hydroxy- und Thiolgruppen von Proteinen reagiert und somit RNasen und andere Enzyme inaktiviert.

### **2.5.1.1 Präparation genomischer DNA**

Für die Präparation der genomischen DNA wurde das NucleoSpin<sup>®</sup> Tissue Kit der Firma Macherey-Nagel (Düren) verwendet. Das frische Gewebepulver wurde den Herstellerangaben entsprechend durch Zusatz von Proteinase K lysiert und auf die Aufreinigungssäulen gegeben. Nach der Bindung der Nukleinsäuren an die Silikamembran der Säule konnte die DNA gewaschen und nach vollständiger Trocknung der Membran mit dem mitgelieferten Puffer eluiert werden.

### **2.5.1.2 Präparation genomischer DNA mit CTAB**

Die in dieser Arbeit angewandte Methode der DNA-Präparation mit Hilfe von CTAB (Cetyltrimethylammoniumbromid) basiert auf einem Protokoll, welches ursprünglich für die Nukleinsäurepräparation aus Pflanzengewebe entwickelt wurde (Doyle & Doyle, 1987; Allen *et al.*, 2006). Das CTAB ermöglicht während der Präparation eine effektive Abtrennung von Polysacchariden, Proteinen und Zelltrümmern. Vor der Präparation wurde wie unter 2.5.1 beschrieben verfahren. Das Pulver von etwa 30 bis 50 mg Gewebe wurde in ein 1,5 ml Reaktionsgefäß überführt und mit 350 µl eines auf 60°C vorgewärmten 2x CTAB Puffers vermischt. Anschließend wurden 25 µl Proteinase K (Konzentration: 20 mg/ml) zugegeben und der gesamte Ansatz erneut vermischt. Danach erfolgte eine 30 minütige Inkubation bei 60°C im Wasserbad. Nach der vollständigen Lyse des Gewebes wurden 350 µl eines Chloroform/Isoamylalkohol-Gemisches (24:1) zugefügt und der gesamte Ansatz für die Dauer von zwei Minuten sorgfältig vermischt. Während der folgenden Zentrifugation (16000xg, 10 Minuten) kam es zu einer Phasenseparation, bei der die obere, wässrige Phase die isolierten Nukleinsäuren enthielt. Nach dem vorsichtigen Überführen der wässrigen Phase in ein neues Reaktionsgefäß wurde eine Nukleinsäurepräzipitation durch Zugabe von 0,7 Volumen Isopropanol durchgeführt. Die ausgefällte DNA wurde durch Zentrifugation sedimentiert, luftgetrocknet und in 300 µl vorgewärmtem (60°C) Wasser (bidestilliert) resuspendiert. Anschließend wurden 20 µl RNase A (Konzentration: 20 mg/ml) zugegeben und der Ansatz vermischt. Dieser Schritt diente dem effektiven Entfernen verbliebener RNA. Nach Zugabe von 300 µl absolutem Ethanol und sorgfältigem Durchmischen der Probe wurde diese auf eine Zentrifugensäule des NucleoSpin<sup>®</sup> Tissue Kits (Macherey-Nagel, Düren) gegeben und nach den

Herstellieranweisungen weiter behandelt. Die Elution der Probe erfolgte mit vorgewärmtem (70°C) TE-Puffer (Tris-EDTA-Puffer).

|                                    |   |
|------------------------------------|---|
| <b>2 x CTAB Isolations-Puffer:</b> | 100 mM Tris/Cl (pH 8,0)<br>1,4 M NaCl<br>20 mM EDTA<br>2 % PVP (Polyvinylpyrrolidon)<br>2 % CTAB (Cetyltrimethylammoniumbromid)<br>0,2 % $\beta$ -Mercaptoethanol (unmittelbar vor<br>Verwendung zugeben) |
| <b>TE-Puffer:</b>                  | 10 mM Tris/Cl (pH 7,6)<br>1 mM EDTA   |

### 2.5.1.3 Präparation von Gesamt-RNA

Die Präparation der Gesamt-RNA erfolgte mit dem peqGOLD Total RNA Kit (PEQLAB, Erlangen), welches nach den Herstellerangaben verwendet wurde. Unmittelbar vor der Präparation wurde das Gewebe wie unter 2.5.1 beschrieben pulverisiert. Nach der Zugabe eines Lysepuffers wurde das Lysat durch eine Shredder-Säule zentrifugiert und auf eine HiBind<sup>®</sup>-Zentrifugensäule geladen. Anschließend erfolgten mehrere Waschschriffe zur Entfernung von Proteinen und Kontaminationen sowie eine vollständige Trocknung der Silikamembran. Die Elution wurde abschließend mit auf 70°C vorgewärmten DEPC-Wasser durchgeführt. Zur langfristigen Lagerung wurde die RNA direkt nach der Präparation in einen -80°C Gefrierschrank verbracht. Die kurzfristige Lagerung erfolgte bei -20°C im Gefrierschrank.

### 2.5.1.4 Isolierung von mRNA aus Gesamt-RNA

Im Rahmen dieser Arbeit wurde die präparierte Gesamt-RNA zur weiteren selektiven Isolierung der mRNA mit Hilfe des nach den Herstellerangaben verwendeten PolyAtract<sup>®</sup> System 1000 (Promega, Mannheim) eingesetzt. Dieses Verfahren basiert auf der Tatsache, dass die mRNA im Gegensatz zu den anderen eukaryotischen RNA-Spezies einen poly-A<sup>+</sup>-Schwanz an ihrem 3'-Ende besitzt. Das verwendete Aufreinigungssystem beinhaltet spezielle, biotinmarkierte Oligo-(dT)-Primer, welche mit dem erwähnten poly-A<sup>+</sup>-Schwanz hybridisieren. Durch die Zugabe von paramagnetischen Partikeln, welche über an sie gekoppeltes Streptavidin spezifisch an die biotinmarkierte mRNA binden können, ist es

möglich, die mRNA in einem magnetischen Feld selektiv zurückzuhalten und von den anderen RNA-Spezies zu trennen. Die Elution der aufgereinigten mRNA erfolgte mit DEPC behandeltem Wasser, welches durch das Absenken der Salzkonzentration zur Ablösung des biotinmarkierten Primers führt.

### **2.5.2 RNA-Aufreinigung mit rDNase-Verdau**

Die wie unter 2.5.1.3 präparierte RNA wurde zur vollständigen Entfernung von genomischer DNA, Proteinen und anderen Kontaminationen mit dem nach Herstellerangaben verwendeten NucleoSpin<sup>®</sup> RNA Clean-up Kit (Macherey-Nagel, Düren) aufgereinigt. Dazu wurde die in DEPC-Wasser gelagerte RNA mit 1/10 Volumen rDNase Enzym-Puffer-Mix versetzt, gemischt und 10 Minuten bei 37°C inkubiert. Anschließend erfolgte die Bindung der RNA an die Silikamembran einer Zentrifugensäule sowie die darauf folgenden Wasch- und Trockenschritte. Für die Elution der RNA wurde auf 70°C vorgewärmtes DEPC-Wasser verwendet. Die Lagerung erfolgte wie unter 2.5.1.3 beschrieben.

### **2.5.3 Phenol/Chloroform Extraktion von Nukleinsäuren**

Zum Entfernen von Proteinkontaminationen aus Nukleinsäurelösungen wurde im Rahmen dieser Arbeit die Methode der Phenol/Chloroform Extraktion angewendet. Die verunreinigte Probe wurde mit einem Volumen eines Phenol-Chloroform-Isoamylalkohol-Gemisches (25:24:1) versetzt und sorgfältig vermischt. Anschließend erfolgte eine 15 minütige Inkubation auf Eis. Danach wurde der Ansatz für 5 Minuten bei 10000xg zentrifugiert und die obere wässrige Phase, welche die Nukleinsäuren beinhaltete, abgenommen und durch eine Ethanolpräzipitation (siehe 2.5.4) weiter aufgereinigt.

### **2.5.4 Ethanolpräzipitation von Nukleinsäuren**

Mit Hilfe der Ethanolpräzipitation (Zeugin *et al.*, 1985) ist es möglich, Verunreinigungen aus Nukleinsäurelösungen zu entfernen und die enthaltenen Nukleinsäuren aufzukonzentrieren. Dafür wurden der wässrigen Nukleinsäurelösung 1/10 des Ausgangsvolumens 3M Natriumacetatpuffer sowie zwei Volumen eiskalter, absoluter Ethanol zugegeben. Nach dem sorgfältigen Vermischen des Ansatzes wurde die Präzipitation für 30 bis 60 Minuten auf Eis durchgeführt. Die ausgefallenen Nukleinsäuren wurden sedimentiert (16000xg, 10 Minuten, 4°C) und nach einem Waschschrift mit 70 % Ethanol

und der anschließenden Trocknung des Sediments in einem geeigneten Puffer resuspendiert.

### 2.5.5 Photometrische Konzentrationsbestimmung

Die Quantifizierung von Nukleinsäuren in Lösungen erfolgte photometrisch mit Hilfe des BioPhotometers der Firma Eppendorf (Hamburg). Dafür wurde die Absorption der Lösung bei einer Wellenlänge von 260 nm detektiert. Bei einer Küvettenschichtdicke von 1 cm und einer Wellenlänge von 260 nm entspricht eine Absorption von 1 OD (optical density):

50 µg/ml dsDNA

37 µg/ml ssDNA

40 µg/ml ssRNA

Die photometrische Messung erlaubt auch Aussagen über die Qualität der untersuchten Probe. Durch Messung der Absorption bei 280 nm, bei der die in der Probe enthaltenen aromatischen Aminosäuren der Proteine absorbieren, kann eine Aussage über die Proteinkontamination der Nukleinsäurelösung getroffen werden. Dafür wird der Quotient  $A_{260}/A_{280}$  gebildet. Dieser sollte bei reiner DNA einen Wert von 1,8 und bei reiner RNA 2,0 erreichen. Des Weiteren wurde die Absorption der Lösung bei 230 nm detektiert. Eine hohe  $A_{230}$  spricht für eine Kontamination der Probe mit Polysacchariden, Kohlenhydraten, Peptiden, Phenolen oder aromatischen Verbindungen. Für eine qualitative Aussage wird der Quotient  $A_{260}/A_{230}$  gebildet, der bei reinen Proben größer 2,0 sein sollte.

### 2.5.6 Reverse Transkription (RT)

Um die isolierte RNA als Matrize für eine Polymerase-Kettenreaktion (polymerase chain reaction, PCR) nutzen zu können, musste diese vorab mit Hilfe des Enzyms „reverse Transkriptase“ in cDNA umgeschrieben werden. Im Rahmen dieser Arbeit wurden sowohl die SuperScript™ III Reverse Transcriptase (Invitrogen, Darmstadt) als auch die Transcriptor Reverse Transcriptase (Roche Applied Science, Mannheim) entsprechend den Herstellerangaben verwendet. Unabhängig vom verwendeten Enzym erfolgte zu Beginn der RT ein Inkubationsschritt von 5-10 Minuten bei 65°C. Die so erreichte Denaturierung der RNA führt zur Auflösung von Sekundärstrukturen und ermöglicht so ein effektives Arbeiten des Enzyms und damit möglichst lange cDNA Transkripte. Der eigentliche

Startpunkt der cDNA-Synthese wird durch den verwendeten Primer festgelegt. Für die vorliegende Arbeit wurden sowohl Oligo-(dT)-Primer als auch genspezifische Primer verwendet. Nach erfolgter RT wurden 1-2 µl des Reaktionsansatzes direkt als Matrize für eine PCR eingesetzt. Der verbleibende Ansatz wurde bis zur weiteren Verwendung bei -20°C gelagert.

## 2.5.7 PCR Techniken

### 2.5.7.1 Primer

Sämtliche in dieser Arbeit verwendeten, genspezifischen und degenerierten Oligonukleotide wurden von Sigma-Aldrich (Taufkirchen) bezogen. Für die Auswahl und Überprüfung der Primer wurden die Programme „PRIMER III“, „DNA CALCULATOR“ und „GENEIOUS“ (Drummond *et al.*, 2011) verwendet (siehe Anhang E). Geeignete Primer sollten folgende Grundsätze erfüllen:

- keine Selbstkomplementarität
- keine Komplementarität zwischen gemeinsam verwendeten Primern
- GC-Gehalt (Guanin + Cytosin) von 40 % bis 60 %
- zueinander passende Annealing-Temperaturen gemeinsam verwendeter Primer (Differenz maximal 5°C)
- Länge zwischen 15 und 30 Nukleotiden

Die Oligonukleotide wurden in lyophilisierter Form geliefert und mussten vor Verwendung in bidestilliertem Wasser resuspendiert werden. Sie wurden bis zur weiteren Verwendung bei -20°C im Gefrierschrank gelagert. Ein Verzeichnis aller verwendeten Primer befindet sich im Anhang dieser Arbeit (F).

### 2.5.7.2 PCR

Bei der Polymerase-Kettenreaktion (PCR) handelt es sich um eine Methode zur *in vitro* Vervielfältigung eines genau definierten DNA-Abschnittes. Für die Durchführung einer PCR benötigt man zumindest eine DNA-Polymerase in einem passenden Puffersystem, zwei Primer, eine DNA- oder cDNA-Matrize (template) sowie genügend Desoxynukleosidtriphosphate (dNTPs), um die neuen DNA-Stränge synthetisieren zu können. Der jeweils zu kopierende Bereich wird dabei durch die entsprechend ausgewählten Oligonukleotid-Primer begrenzt. Eine typische PCR beginnt mit einem


einmaligen, initialen Denaturierungsschritt, im Verlaufe dessen der Reaktionsansatz auf 94°C bis 95°C erwärmt und somit die doppelsträngige DNA-Matrize in die beiden Einzelstränge getrennt wird. Anschließend folgt ein zyklischer Ablauf von drei verschiedenen Temperaturstufen, die der Vervielfältigung des ausgewählten DNA-Bereiches dienen. Ein Zyklus wiederum beginnt mit einer kurzen Denaturierung bei ~94°C. Anschließend erfolgt ein als Annealing bezeichneter Schritt, der der Anlagerung der Oligonukleotid-Primer an die DNA-Matrize dient. Die in diesem Schritt verwendete Temperatur richtet sich dabei nach der Schmelztemperatur ( $T_m$ ) der verwendeten Primer und muss für jede neue PCR optimiert werden. Der dritte als Elongation bezeichnete Schritt beinhaltet die Synthese des zur Matrize komplementären Stranges beginnend mit dem freien 3'-OH-Ende des Primers. Dabei fügt die thermostabile DNA-Polymerase bei einer Temperatur zwischen 68°C und 72°C für die gewählte Zeitdauer die komplementären dNTPs an den neuen Strang an. Die zeitliche Dauer dieses Schrittes richtet sich nach der Größe des zu vervielfältigenden Abschnittes. Als Richtwert wird eine Minute pro 1000 zu amplifizierenden Basenpaaren angenommen. Der beschriebene Zyklus wird 25 bis 35 mal wiederholt und führt zu einer exponentiellen Vermehrung des gewünschten Bereiches. Abschließend erfolgt ein deutlich längerer finaler Elongationsschritt, welcher zur Fertigstellung aller nicht vollständig synthetisierten DNA-Fragmente dient.

Im Verlauf dieser Arbeit wurden Polymerasen mit unterschiedlichen Eigenschaften verwendet. Der DreamTaq™ Green PCR Master Mix (Fermentas, St. Leon-Rot) beinhaltet alle für die Reaktion erforderlichen Komponenten mit Ausnahme der Primer und der DNA-Matrize. Zusätzlich ist auch ein Farbstoff zugesetzt, der ein direktes Auftragen des Reaktionsansatzes auf ein Agarosegel ermöglicht. Bei der in dieser Mischung enthaltenen Polymerase handelt es sich um die thermostabile DNA-Polymerase des Bakteriums *Thermophilus aquaticus* (Taq-Polymerase). Diese Polymerase besitzt unter anderem die Eigenschaft, ein Desoxyadenin an das 3'-Ende aller synthetisierten DNA-Stränge anzufügen. Diese Überhänge sind essentiell für das ebenfalls im Rahmen dieser Arbeit durchgeführte TA-Klonierungsverfahren (siehe 2.5.9).

Ein typischer 25 µl Reaktionsansatz unter Verwendung dieser Polymerase war wie folgt zusammengesetzt:

|              |                                     |
|--------------|-------------------------------------|
| 12,5 $\mu$ l | DreamTaq™ Green PCR Master Mix      |
| 8,5 $\mu$ l  | H <sub>2</sub> O bidest             |
| 1 $\mu$ l    | <i>forward</i> -Primer (10 $\mu$ M) |
| 1 $\mu$ l    | <i>reverse</i> -Primer (10 $\mu$ M) |
| 2 $\mu$ l    | cDNA Matrize                        |

Alle Reaktionen wurden auf Eis angesetzt, gut gemischt, kurz zentrifugiert und in den bereits entsprechend vorgeheizten Thermocycler (T Gradient Thermocycler 96; Biometra GmbH, Göttingen) platziert. Ein typisches PCR-Ablaufschema sah wie folgt aus:

|                        |      |          |  |
|------------------------|------|----------|--|
| Initiale Denaturierung | 95°C | 3 min    |  |
| Denaturierung          | 94°C | 30 sec   |  35 x |
| Annealing              | 59°C | 30 sec   |  |
| Elongation             | 72°C | 3 min    |  |
| Terminale Elongation   | 72°C | 7 min    |  |
| Kühlung                | 4°C  | $\infty$ |  |

Dieses Standardprogramm wurde entsprechend den speziellen Anforderungen der durchgeführten PCR in unterschiedlicher Weise verändert. Bei der erstmaligen Verwendung neuer Primer wurden zum Beispiel Reaktionen mit variierenden Annealingtemperaturen durchgeführt, um so die optimalen Bedingungen für die jeweiligen Primerkombinationen zu ermitteln. Da es sich bei dem verwendeten Thermocycler um ein Gradientenmodell handelte, konnten die Reaktionen mit den verschiedenen Temperaturen gleichzeitig vollzogen werden.

Eine weitere verwendete Polymerase war die *Taq* DNA Polymerase von Invitrogen (Darmstadt). Bei diesem Enzym lagen die einzelnen Komponenten des Reaktionsansatzes getrennt vor, so dass hier weitergehende Variationen des Ansatzes bezüglich Enzym-, MgCl<sub>2</sub>- und dNTP-Konzentration möglich waren. Der Ablauf der PCR basierte auch hier auf dem oben beschriebenen Standardprogramm.

### 2.5.7.3 Nested PCR

Bei der nested PCR (verschachtelte PCR) handelt es sich um zwei aufeinander folgende Reaktionen mit unterschiedlichen Primer-Paaren. Diese PCR-Technik wird verwendet, um

die Spezifität der Reaktion zu erhöhen und die Bildung unerwünschter Produkte zu minimieren. In der ersten Reaktion wird ein DNA-Fragment mit einer bestimmten Kombination aus *forward*- und *reverse*-Primer amplifiziert. Ein Aliquot dieser Reaktion wird dann als Matrize für die zweite PCR eingesetzt. In dieser Reaktion werden Primer verwendet die innerhalb des durch die Primer der ersten Reaktion begrenzten Bereiches liegen. Da es höchst unwahrscheinlich ist, dass unerwünschte Sekundärprodukte Bindungsstellen für beide Primer-Paare enthalten, kann man mit Hilfe dieser PCR-Technik Fragmente mit hoher Spezifität amplifizieren.

#### **2.5.7.4 Touchdown PCR**

Die Methode der touchdown oder stepdown PCR dient der Erhöhung der Spezifität der Reaktion durch die Forcierung der akkuraten Primerbindung. Das Prinzip dieser Methode ist die schrittweise Annäherung an die optimale Annealingtemperatur. Die Obergrenze der möglichen Annealingtemperatur wird dabei durch die Schmelztemperatur ( $T_m$ ) der Primer festgelegt. Liegt die Annealingtemperatur nur knapp unterhalb der  $T_m$  sind nur hochspezifische Primerbindungen an die Matrize möglich. In der praktischen Anwendung wurde eine initiale Annealingtemperatur gewählt, die  $5^\circ\text{C}$  bis  $10^\circ\text{C}$  oberhalb der  $T_m$  der verwendeten Primer lag. Mit jedem fortschreitenden PCR-Zyklus wurde nun sukzessive die Annealingtemperatur verringert, bis sie schließlich knapp unterhalb der  $T_m$  der Primer lag. Durch diese Verfahrensweise wurde erreicht, dass die ersten amplifizierten Fragmente mit hochspezifischen Primerbindungen entstehen. Im Verlauf der PCR wurde die Annealingtemperatur auf einen Wert von etwa  $5^\circ\text{C}$  unterhalb der  $T_m$  gesenkt. Bei diesen Temperaturen könnte es zwar auch zu unspezifischeren Primerbindungen kommen, jedoch wird durch den exponentiellen Charakter der PCR vorwiegend das spezifische Anfangsfragment weiter amplifiziert. Die Nomenklatur dieser Methode ergibt sich aus den Schrittweiten der Annealingtemperaturabnahme. Wenn die Verringerung in großen Schritten von etwa  $1^\circ\text{C}$  vorgenommen wird, spricht man von einer stepdown-, bei kleineren Schritten von einer touchdown-PCR.


#### **2.5.7.5 Band-stab PCR**

Die Methode der Band-stab PCR (Bjourson *et al.*, 1992) verfolgt einen anderen Ansatz, um spezifischere Fragmente amplifizieren zu können. Bei der Durchführung von Polymerase-Kettenreaktionen kommt es häufig nicht nur zur Vervielfältigung der gewünschten spezifischen sondern auch von verschiedenen unspezifischen DNA-Fragmenten. Für eine Band-stab PCR wird anfänglich eine PCR mit einem bestimmten Primer-Paar

durchgeführt. Der Ansatz wird in einem Agarosegel aufgetrennt und das Bandenmuster unter ultravioletter Beleuchtung sichtbar gemacht (siehe 2.5.8.1). Jetzt wird ein der ersten PCR entsprechender Ansatz vorbereitet, der jedoch statt der Matrize nur bidestilliertes H<sub>2</sub>O enthält. Die Primerkombination sollte bei diesem Ansatz der des ersten Ansatzes entsprechen. Nun wird die Oberfläche des Gels mit einem Papiertuch getrocknet und die gewünschte Bande im Gel lokalisiert. Unter Zuhilfenahme einer sterilen Stahlkanüle wird die entsprechende Bande mehrmals punktiert und die Nadel sofort in einem frischen PCR-Ansatz leicht geschwenkt. Der so mit einer DNA-Matrize versehene PCR-Ansatz kann nun wieder in einem Thermocycler platziert und nach erfolgter PCR erneut auf einem Agarosegel aufgetrennt werden. Mit Hilfe dieser Methode ist es möglich einzelne Banden aus einer Mischung von DNA-Fragmenten gezielt als Matrize für eine folgende PCR zu verwenden und somit unerwünschte Nebenprodukte zu minimieren.

#### 2.5.7.6 Klon PCR

Mit Hilfe der Klon PCR wurde nach einer Klonierungsreaktion überprüft, ob der in den Bakterien replizierte Plasmidvektor ein Integrat der gewünschten Länge beinhaltet. Unter Verwendung des DreamTaq™ Green PCR Master Mix wurden PCR-Ansätze (12 µl) mit vektorspezifischen Primern (siehe Anhang F.1) hergestellt. Mit einer sterilen Pipettenspitze wurde eine Bakterienkolonie von einer Plattenkultur (siehe 2.4.2) abgestreift und im PCR-Ansatz auf- und abpipettiert. Die Pipettenspitze wurde dann in ein 15 ml Reaktionsgefäß überführt und für das Animpfen einer Flüssigkultur (siehe 2.4.3) aufbewahrt. Die PCR-Reaktion wurde grundsätzlich wie unter 2.5.7.2 beschrieben durchgeführt. Das Ablaufschema wurde jedoch an die Erfordernisse der Klon-PCR angepasst:

|                        |      |           |   |
|------------------------|------|-----------|---|
| Initiale Denaturierung | 95°C | 3 min     |   |
| Denaturierung          | 94°C | 30 sec    | <br>24 x |
| Annealing              | 53°C | 30 sec    |   |
| Elongation             | 72°C | 1 min/kBp |   |
| Terminale Elongation   | 72°C | 7 min     |   |
| Kühlung                | 4°C  | ∞         |   |

### 2.5.7.7 GeneRacer

Das GeneRacer™ Kit von Invitrogen (Darmstadt) ermöglicht es vollständige cDNAs zu amplifizieren (siehe Abb. 11). Dabei steht der Begriff „RACE“ für die „rapid amplification of cDNA ends“. Im Rahmen dieser Arbeit wurde das GeneRacer™ Kit verwendet, um die unbekannt 5′-Bereiche der untersuchten cDNAs zu analysieren. Mittels des Enzyms „CIP“ (calf intestinal phosphatase) werden dabei zunächst alle nicht vollständigen mRNAs sowie alle anderen RNAs ohne Cap-Struktur am 5′-Ende dephosphoryliert und somit von der folgenden Ligationsreaktion ausgeschlossen. Danach werden die Cap-Strukturen von den intakten, vollständigen mRNAs durch das Enzym „TAP“ (tabacco acid pyrophosphatase) entfernt.

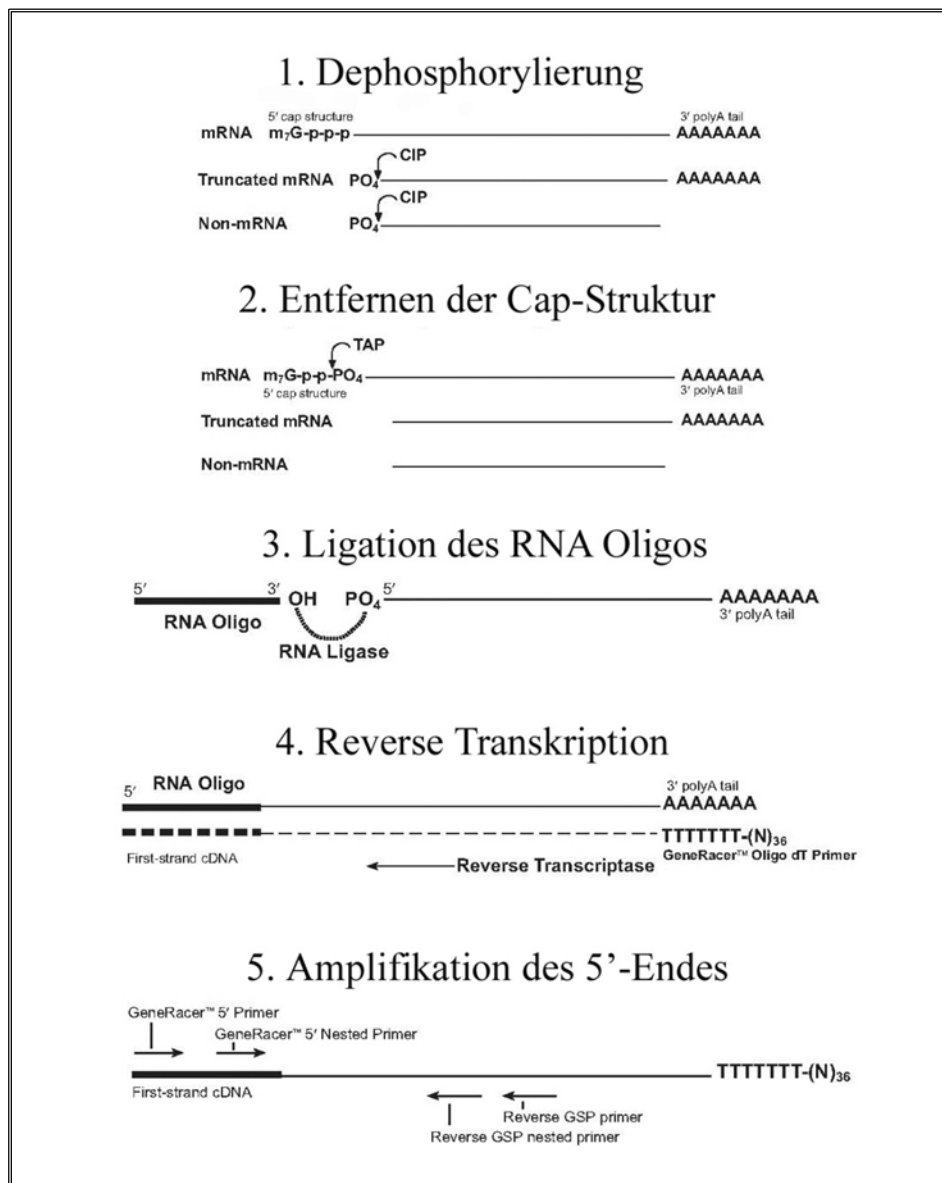


Abb. 11: Amplifikation der 5′-Enden der untersuchten Gene mit dem GeneRacer™ Kit (nach Invitrogen, Darmstadt)

An das 5'-Ende der so vorbereiteten mRNAs kann nun mit Hilfe der T4 RNA Ligase ein spezielles RNA-Oligonukleotid ligiert werden. Nach der nun folgenden reversen Transkription (siehe 2.5.6) der mRNA mittels eines genspezifischen *reverse*-Primers folgt eine nested-PCR zur Amplifikation der unbekanntenen Fragmente des 5'-Endes. Die dabei verwendeten *forward*-Primer (im GeneRacer™ Kit enthalten) binden an eine definierte Sequenz des ligierten RNA-Oligonukleotids, während die genspezifischen *reverse*-Primer innerhalb des bereits bekannten Genbereiches liegen. Zur weiteren Erhöhung der Spezifität des amplifizierten Fragmentes konnte eine zweite nested PCR unter Verwendung eines zusätzlichen Primersatzes durchgeführt werden

#### **2.5.7.8 PCR-Aufreinigung**

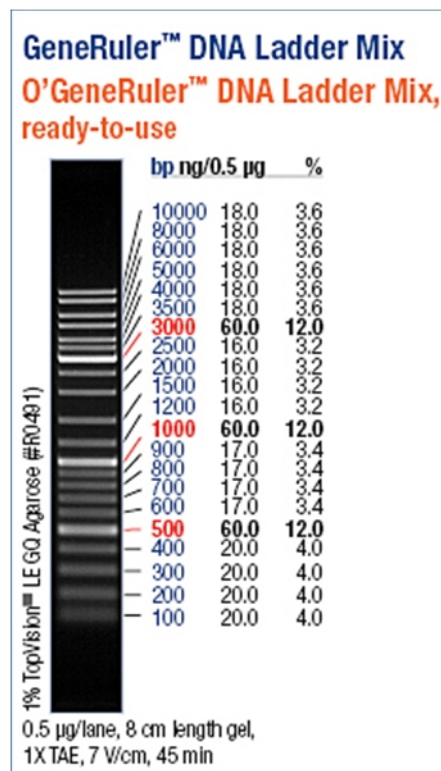
Um die generierten PCR-Produkte aus dem PCR-Ansatz aufzureinigen und somit die weiteren Komponenten des Ansatzes zu entfernen, wurde eine PCR-Aufreinigung mit dem GeneJET™ PCR Purification Kit (Fermentas, St. Leon-Rot) durchgeführt. Diese Methode basiert ebenso wie die DNA-Extraktion aus Agarosegelen (2.5.8.3) auf der reversiblen Bindung der DNA an die Silikamembran einer Zentrifugensäule. Der PCR-Ansatz wurde den Herstellerangaben entsprechend mit dem Bindepuffer vermischt und auf die Säule gegeben. Unter den durch den Puffer erzeugten Hochsalzbedingungen konnte die DNA an die Membran gebunden werden. Im Folgenden wurden zwei Waschschriffe durchgeführt und die DNA nach vollständiger Trocknung der Membran mit 20 µl des auf 70°C vorgewärmten Elutionspuffers eluiert.

### **2.5.8 Gelelektrophoretische Techniken**

#### **2.5.8.1 Agarose-Gelelektrophorese von DNA**

Die Agarose-Gelelektrophorese wurde verwendet, um DNA-Fragmente entsprechend ihrer Größe aufzutrennen. Dabei wandert die durch das Zucker-Phosphat-Rückgrat negativ geladene DNA im elektrischen Feld zur Anode. Die Wanderungsgeschwindigkeit richtet sich nach der Höhe der angelegten Spannung, der durch die Agarosekonzentration bestimmten Porengröße des Gels sowie nach der Größe der aufzutrennenden DNA-Fragmente. Für diese Arbeit wurden horizontale Gele verwendet, deren Agarosekonzentration je nach erforderlichem Auftrennungsbereich zwischen 0,8 % und 2,0 % lag. Zur späteren Detektion der DNA unter ultravioletter Beleuchtung (UV-Transilluminator der Wellenlänge 312 nm; Vilber Lourmat, Eberhardzell) wurden dem Agarosegel 0,5 µl Ethidiumbromid pro Milliliter Gelvolumen zugesetzt. Bei Ethidiumbromid handelt es sich

um einen roten Phenanthridin-Farbstoff, der in Nukleinsäuren interkalieren kann und somit einen qualitativen und orientierenden quantitativen Nachweis von DNA im Agarosegel ermöglicht. Die Leuchtintensität des Ethidiumbromids ist dabei proportional zur Nukleinsäurekonzentration. Zu Beginn der Gelelektrophorese wurde das entsprechende Gel in der Kammer platziert und mit Elektrophoresepuffer bedeckt. Danach wurden die mit Beladungspuffer vermischten Proben in die Geltaschen pipettiert. In die äußersten Taschen des Gels wurde zudem ein Größenstandard aufgetragen (*GeneRuler™ DNA Ladder Mix*; Fermentas, St. Leon-Rot), mit dessen Hilfe später die Größe der aufgetrennten Fragmente ermittelt werden konnte. Die so vorbereitete Gelkammer (Biotech Fischer, Reiskirchen) wurde nun an eine konstante Spannungsquelle angeschlossen (6 V/cm Elektrodenabstand) und die Auftrennung der DNA-Fragmente durchgeführt. Abschließend wurden die Agarosegele mittels eines Geldokumentationssystems (PEQLAB, Erlangen) mit angeschlossener CCD-Kamera (Sony, Berlin) dokumentiert.



**Abb. 12:** Verwendeter Größenstandard GeneRuler™ DNA Ladder Mix (Fermentas, St. Leon-Roth)

|                             |  |
|-----------------------------|--|
| <b>50 x TAE-Puffer:</b>     | 2 M Tris (Tris(hydroxymethyl)-aminomethan)<br>100 mM EDTA (Ethylendiamintetraacetat)<br>57,1 ml Eisessig<br>mit H <sub>2</sub> O dest. auf 1 Liter auffüllen |
| <b>1 x TAE-Puffer:</b>      | 20 ml 50 x TAE-Puffer<br>mit H <sub>2</sub> O dest. auf 1 Liter auffüllen<br>als Laufpuffer für die Elektrophorese verwenden                                 |
| <b>10 x TBE-Puffer:</b>     | 0,9 M Tris/Cl<br>0,9 M Borsäure<br>20 mM EDTA<br>mit H <sub>2</sub> O dest. auf 1 Liter auffüllen  |
| <b>1 x TBE Puffer:</b>      | 100 ml 10 x TBE-Puffer<br>mit H <sub>2</sub> O dest. auf 1 Liter auffüllen<br>als Laufpuffer für die Elektrophorese verwenden                                |
| <b>Ethidiumbromid:</b>      | Stammlösung mit 10 mg/ml   |
| <b>6 x Beladungspuffer:</b> | 10 mM Tris<br>0,03 % Bromphenolblau<br>0,03 % Xylencyanol FF<br>60 % Glyzerin<br>60 mM EDTA<br>(Fermentas, St. Leon-Roth)                                    |

### 2.5.8.2 Agarose-Gelelektrophorese von RNA

Zur RNA-Qualitätsbeurteilung wurde nach erfolgter Präparation eine Agarose-Gelelektrophorese durchgeführt. Im Gegensatz zur Gelelektrophorese von DNA wurde hier dem Probenpuffer sowie dem Gel Formaldehyd zugesetzt, welches die Ausbildung von Sekundärstrukturen und die Degradation der RNA durch ubiquitäre RNasen verhindern sollte. Die RNA-Probe wurde mit 1/10 Probenpuffer vermischt und für 3 Minuten auf 70°C erhitzt. Danach erfolgte eine zweiminütige Inkubation im Eiswasser. Nachdem die Probe in die Taschen des Gels pipettiert wurde, erfolgte die Elektrophorese bei einer

konstanten Spannung von 100 V. Die verwendete horizontale Minigelkammer (Biotech Fischer, Reiskirchen) wurde für die Dauer der Elektrophorese auf Eis platziert, um das Schmelzen des Gels durch Überhitzen zu verhindern. Durch das im Probenpuffer enthaltene Ethidiumbromid konnte die RNA nach erfolgter Elektrophorese unter UV-Beleuchtung sichtbar gemacht werden. Die Dokumentation erfolgte mittels eines Geldokumentationssystems (PEQLAB, Erlangen) mit angeschlossener CCD-Kamera (Sony, Berlin).

|                           |  |
|---------------------------|--|
| <b>10 x MOPS-Puffer:</b>  | 200 mM MOPS (3-(N-Morpholino)-2-Hydroxypropansulfonsäure-Natriumsalz)<br>50 mM Natrium-Acetat<br>10 mM EDTA (Ethyldiamintetraacetat)<br>pH 7<br>lichtgeschützte Lagerung |
| <b>1 x MOPS-Puffer:</b>   | 100 ml 10 x MOPS-Puffer<br>900 ml H <sub>2</sub> O bidest.<br>als Laufpuffer für die Elektrophorese verwenden  |
| <b>10 x Probenpuffer:</b> | 50 % Glycerin<br>0,4 % Bromphenolblau<br>1 mM EDTA<br>25 % Formaldehyd<br>0,5 µg/ml Ethidiumbromid   |
| <b>RNA-Minigel:</b>       | 4 ml 10 x MOPS-Puffer<br>30 ml H <sub>2</sub> O bidest.<br>0,4 g Agarose<br>Gemisch aufkochen und auf 50°C abkühlen<br>6 ml Formaldehyd zufügen<br>Gel sofort gießen     |

### 2.5.8.3 Extraktion von DNA aus Agarosegelen

Die gewünschten DNA-Fragmente wurden unter Verwendung des GeneJET™ Gel Extraction Kit (Fermentas, St. Leon-Rot) nach der erfolgten Elektrophorese wieder aus dem Agarosegel extrahiert. Dafür wurde die gewünschte Bande unter UV-Beleuchtung betrachtet und mittels eines sterilen Skalpellens aus dem Gel herausgeschnitten. Das Gelstück wurde mit dem im Kit enthaltenen Bindepuffer vermischt und für 5 bis 10 Minuten bei 55°C bis 65°C inkubiert. Im Folgenden wurde das so verflüssigte Gelstück auf eine Aufreinigungssäule gegeben. Die enthaltene DNA wurde nach Bindung an die Silikamembran der Säule zur Entfernung von Kontaminationen zweifach gewaschen. Abschließend erfolgte die Elution der aufgereinigten DNA mit 20 µl des auf 70°C vorgewärmten Elutionspuffers.

### 2.5.9 Klonierung von PCR-Produkten

Im Rahmen dieser Arbeit wurde die Methode der Klonierung von PCR-Produkten verwendet um DNA-Fragmente nach vorheriger gelelektrophoretischer Auftrennung und anschließender Gelextraktion zu vervielfältigen. Die DNA wurde dafür in einen Vektor ligiert und in bakterielle Wirtszellen transformiert. Bei jeder Teilung der Bakterien wird der Vektor mit dem gewünschten Fragment kopiert und an die Tochterzellen weitergegeben. Diese Methode ermöglicht somit das Kopieren der DNA-Fragmente in hoher Anzahl und sehr reiner Form. Für die vorliegende Arbeit wurde zum einen das StrataClone™ PCR Cloning Kit (Agilent, Böblingen) und zum anderen das TOPO® XL PCR Cloning Kit (Invitrogen, Darmstadt) verwendet, wobei letzteres für besonders lange DNA-Fragmente zum Einsatz kam. Beide verwendeten Produkte basieren auf dem Prinzip der TA-Klonierung, welche keine enzymatische Restriktion der eingesetzten DNA-Fragmente benötigt. Durch die während der PCR-Amplifikation verwendete *Taq*-Polymerase wird ein Desoxyadenin an das 3'-Ende aller synthetisierten DNA-Stränge angefügt. Die linear vorliegenden Vektorarme besitzen an jeweils einem Ende ein 3'-Desoxyuracil und sind mit einer kovalent gebundenen Topoisomerase I beladen. Die Desoxyadeninüberhänge des PCR-Produktes sind somit kompatibel zu den Vektorarmen und können mit den dort überhängenden Desoxyuracilen paaren. Die gebundene Topoisomerase I bewirkt dann nach der Basenpaarung eine Ligation des DNA-Fragmentes mit den beiden Vektorarmen (siehe Abb. 13). Das resultierende lineare DNA-Konstrukt wurde dann durch einen Hitzeschock in chemisch kompetente Bakterien transformiert. Dafür wurden die Bakterien mit der Ligationsreaktion vermischt und für

20 Minuten auf Eis inkubiert. Im Anschluss wurde der Transformationsansatz für 45 Sekunden bei 42°C im Wasserbad positioniert.

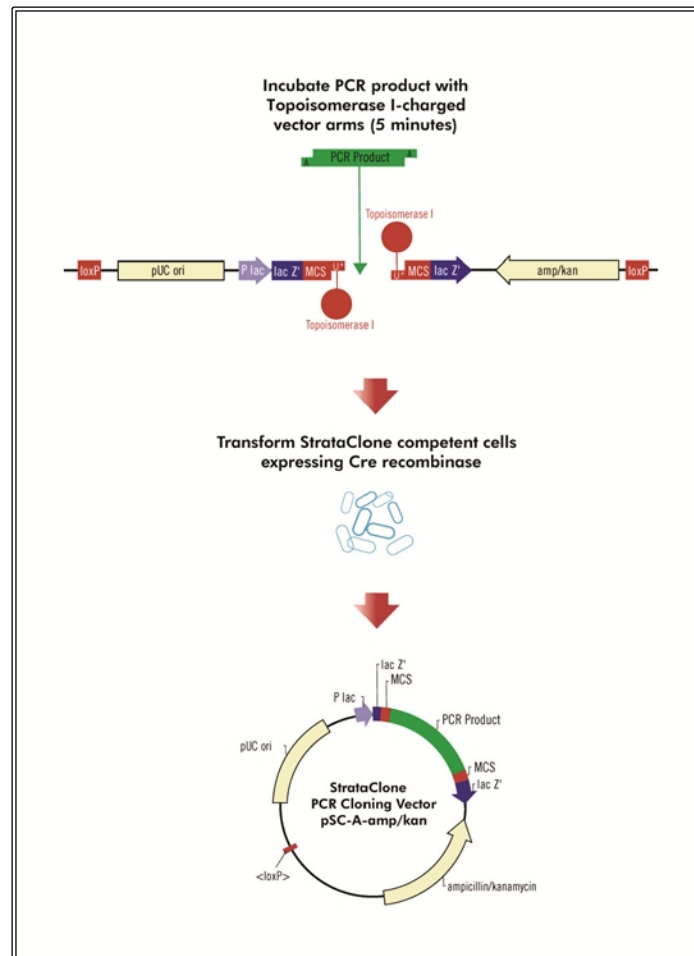


Abb. 13: Übersicht der StrataClone™ Klonierungsmethode (nach Agilent, Böblingen)

Danach erfolgte nochmals eine zweiminütige Inkubation auf Eis. Die verwendeten Bakterienstämme (siehe 2.4.5.1) exprimieren transient eine Cre Rekombinase. Dieses Enzym vermittelt eine Rekombinationsreaktion zwischen den am noch freien Ende der beiden Vektorarme vorhandenen *loxP*-Motiven und führt somit zur Zirkularisation des DNA-Konstruktes. Nach der erfolgten Transformation wurden die Bakterien mit 250 µl vorgewärmten SOC-Medium vermischt und für 60 Minuten bei 37°C geschüttelt. Im Anschluss wurde der Ansatz auf Agarplatten mit einem der Resistenz des Vektors entsprechenden Antibiotikum ausplattiert und über Nacht bei 37°C im Brutschrank inkubiert. Sämtliche Klonierungsreaktionen wurden nach den Angaben der Hersteller durchgeführt. Die eingesetzten Mengen der Komponenten für die Ligationsreaktionen mit dem StrataClone™ PCR Cloning Kit wurden jedoch halbiert.

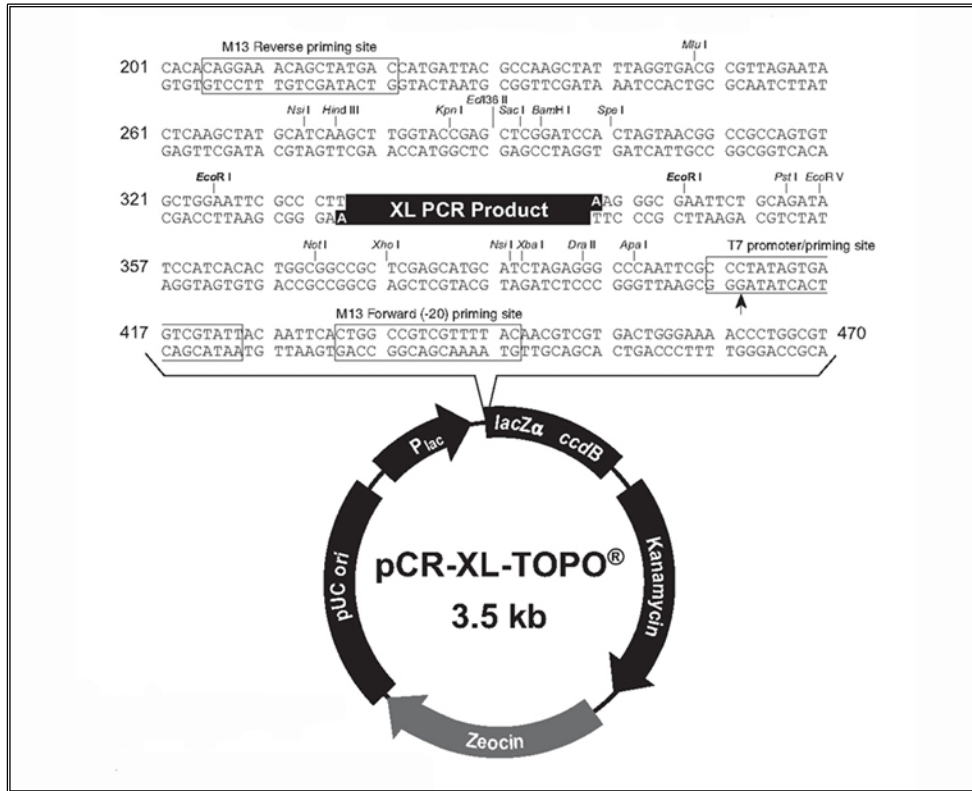


Abb. 14: pCR-XL-TOPO® Klonierungsvektor (nach Invitrogen, Darmstadt)

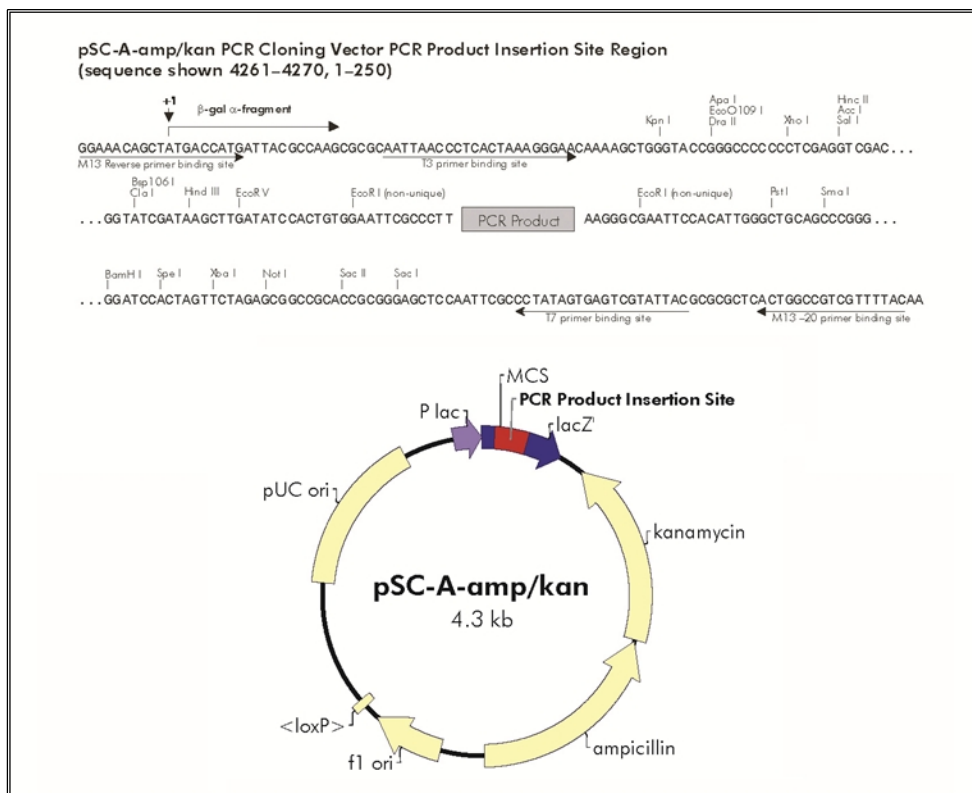


Abb. 15: StrataClone™ PCR Cloning Vector pSC-A-amp/kan (nach Agilent, Böblingen)

### 2.5.9.1 Blau-Weiß-Selektion

Nach erfolgter Ligation und Transformation wurden die Bakterien auf antibiotikahaltigen Agarplatten ausgestrichen und über Nacht bei 37°C im Brutschrank inkubiert (siehe 2.5.9). Anschließend mussten die Bakterienklone identifiziert werden, die den Vektor aufgenommen und bei denen das PCR-Produkt bzw. das Insert erfolgreich in den Vektor ligiert werden konnte. Die Aufnahme des Vektors verlieh den Bakterien eine bestimmte, vektorabhängige Antibiotikaresistenz. Im Falle des StrataClone™ PCR Cloning Vector pSC-A-amp/kan (Agilent, Böblingen) wurde ein Wachstum in Gegenwart von Kanamycin sowie Ampicillin und im Falle des pCR-XL-TOPO® Vektors (Invitrogen, Darmstadt) in Gegenwart von Kanamycin und Zeocin ermöglicht. Aus diesem Grund waren nur erfolgreich mit dem Vektor transformierte Bakterienklone auf den verwendeten Agarplatten lebensfähig. Die Blau-Weiß-Selektion ermöglichte die Identifizierung der Bakterienklone mit den inserttragenden Vektoren. Alle in dieser Arbeit verwendeten Bakterienstämme (siehe 2.4.5.1) wiesen eine Mutation im lacZ-Gen auf, welches bei ihnen nicht für das  $\alpha$ -Fragment der  $\beta$ -Galaktosidase kodierte. Der fehlende Anteil des Enzyms wurde erst durch den transformierten Plasmidvektor zur Verfügung gestellt ( $\alpha$ -Komplementierung). Allerdings war dieser Bereich auf dem Vektor nur ablesbar, wenn keine erfolgreiche Ligation eines Inserts in den Vektor stattgefunden hatte. Somit wiesen nur vektortragende Bakterienklone ohne das entsprechende Insert eine  $\beta$ -Galaktosidaseaktivität auf. Der Nachweis der vorhandenen bzw. fehlenden  $\beta$ -Galaktosidaseaktivität erfolgt durch den Zusatz des Substrates X-Gal (5-Brom-4-Chlor-3-Indoyl- $\beta$ -D-Galactopyranosid) zu den verwendeten Agarplatten. Wenn die  $\beta$ -Galaktosidase exprimiert wurde, war sie in der Lage das X-Gal zu Galaktose und 5-Brom-4-Chlor-3-Hydroxyindol zu hydrolysieren. Letzteres wurde durch den Luftsauerstoff zu einem Farbstoff oxidiert welcher die entsprechenden Bakterienkolonien tiefblau erscheinen ließ. Somit konnten reinweiße Bakterienkolonien als erfolgreich transformiert und inserttragend identifiziert werden. Diese Kolonien wurden mit einer sterilen Pipettenspitze aufgenommen und zum Ansetzen einer Flüssigkultur (siehe 2.4.3) verwendet.

### 2.5.9.2 Selektion mit Hilfe des letalen *E. coli* Gens

Im Falle des pCR-XL-TOPO® Vektors (Invitrogen, Darmstadt) wurde die Identifizierung der für den Vektor und das entsprechende Insert positiven Bakterienklone mit Hilfe des letalen *E. coli* Gens (*ccdB*) durchgeführt. Der genannte Vektor trägt das *ccdB* Gen, welches hier mit dem C-Terminus des lacZ $\alpha$ -Fragmentes fusioniert ist. Durch die Expression des Genproduktes wird die bakterielle DNA-Gyrase inhibiert und das

Wachstum der Bakterien verhindert. Die Ligation eines PCR-Produktes bewirkt hier die Zerstörung des Leserasters sowohl für das *lacZ $\alpha$* - als auch für das *ccdB*-Gen. Somit konnten die auf den Agarplatten wachsenden (weißen) Bakterienkolonien als erfolgreich transformiert und inserttragend ermittelt werden.

### 2.5.9.3 Präparation von Plasmid-DNA

Für die Präparation von Plasmid-DNA aus Bakterien wurde im Rahmen dieser Arbeit das GeneJET™ Plasmid Miniprep Kit (Fermentas, St. Leon-Roth) den Herstellerangaben entsprechend verwendet. Um die rekombinanten Plasmidvektoren aus den Bakterien einer Flüssigkultur zu isolieren, wurde der Ansatz zu Beginn für 5 Minuten bei 6000xg zentrifugiert. Das Bakterienpellet wurde in einem RNase-haltigen Puffer resuspendiert und in ein 1,5 ml Reaktionsgefäß überführt. Durch Zugabe einer Lösung wurde eine alkalische Lyse durchgeführt und der Ansatz danach durch eine weitere Lösung wieder neutralisiert. Das Lysat wurde zentrifugiert (16000xg, 5 Minuten) und somit Zelltrümmer und chromosomale DNA pelletiert. Der klare Überstand wurde auf eine Zentrifugensäule geladen, an deren Silikamembran die Plasmid-DNA reversibel binden konnte. Nach zwei Waschschrritten wurde die Plasmid-DNA durch Zugabe von vorgewärmtem (70°C) Elutionspuffer eluiert. Die isolierte Plasmid-DNA wurde wie unter 2.5.5 beschrieben einer photometrischen Konzentrationsbestimmung unterzogen.

### 2.5.9.4 Restriktionsverdau von Plasmid-DNA

Als Alternative zur Klon PCR (siehe 2.5.7.6) konnte die Integratgröße eines Plasmidvektors nach der Klonierung durch einen Restriktionsverdau mit folgender Gelelektrophorese ermittelt werden. Die in dieser Arbeit verwendeten Plasmidvektoren verfügen über zwei das Integrat flankierende Schnittstellen, die durch die Restriktionsendonuklease *EcoRI* erkannt werden. Bei dem genannten Enzym handelt es sich um eine Restriktionsendonuklease des Typs II. Diese Enzyme sind in der Lage eine 4-8 Basenpaare lange palindromische Erkennungssequenz zu identifizieren und die DNA an genau dieser Stelle zu schneiden. Das in dieser Arbeit verwendete Enzym wurde von Fermentas (St. Leon-Roth) bezogen und nach den Herstellerangaben verwendet. Nach der enzymatischen Restriktion wurden sowohl die verdauten Ansätze als auch das unverdaute Plasmid auf ein Agarosegel aufgetragen und die Gelelektrophorese wie unter 2.5.8.1 beschrieben durchgeführt.

## 2.6 Sequenzanalyse

### 2.6.1 DNA-Sequenzierung

Die im Rahmen dieser Arbeit generierten DNA-Fragmente wurden zur Sequenzierung an die StarSEQ<sup>®</sup> GmbH (Mainz) übergeben. Dafür wurden vorab 400 bis 700 ng Plasmid-DNA bzw. 50 bis 200 ng PCR-Produkt mit einem Mikroliter einer 10 µM Primerlösung (genspezifisch oder vektorspezifisch) vermischt. Anschließend wurde die Probe durch die Firma StarSEQ<sup>®</sup> mit Hilfe eines ABI 3730 DNA Analyzer nach dem Sanger-Verfahren sequenziert. Dieses etablierte Sequenzierverfahren wird als Didesoxymethode oder Kettenabbruch-Synthese bezeichnet. Es beruht auf dem zufälligen Einbau von dem Sequenzieransatz zugefügten Didesoxynukleotiden. Diese machen aufgrund der fehlenden OH-Gruppe am 3'-C-Atom der Ribose eine weitere Verlängerung des in Synthese befindlichen DNA-Stranges unmöglich. Durch dieses Verfahren entstehen dann DNA-Fragmente mit unterschiedlicher Länge, die nach Auftrennung in einem Gel Rückschlüsse auf die Basenabfolge des eingesetzten DNA-Fragmentes erlauben. Nach erfolgter Sequenzierung wurden die Rohdaten durch die Firma StarSEQ<sup>®</sup> in elektronischer Form (Elektropherogramme und FASTA-Sequenzdateien) zur Verfügung gestellt. Die erhaltenen Daten wurden anschließend mit verschiedenen Softwarelösungen (siehe 2.6.2 sowie Anhang E) analysiert und verarbeitet. Auf die verwendeten Programme und Methoden wird an den entsprechenden Stellen im Text weiter eingegangen.

### 2.6.2 Sequenzanalyse

Die erhaltenen Elektropherogramme der Sequenzierungen wurden mit dem Programm „FINCHTV“ betrachtet und hinsichtlich Qualität und Lesbarkeit initial beurteilt. Bei Sequenzierungen mit vektorspezifischen Primern wurden verbleibende Sequenzanteile der Plasmidvektoren entfernt und die Deszendenz der Sequenz durch den Abgleich mit den Internetdatenbanken des „NCBI“ (National Center for Biotechnology Information) identifiziert. Die DNA-Fragmente, die als Teil eines Hämocyanin-Gens bestätigt werden konnten, wurden im Folgenden für die Rekonstruktion der analysierten Gene weiter verwendet. Die Übersetzung der Nukleotid- in Aminosäuresequenzen erfolgte sowohl mit dem Internetprogramm „FR33.NET TRANSLATOR“ als auch mit dem Softwarepaket „GENEIOUS“ (Drummond *et al.*, 2011). Die Assemblierung und Überprüfung der Mega-Hämocyanin-Gene aus den einzelnen DNA-Fragmenten erfolgte unter Verwendung des

Programms „GENEIOUS“ (Drummond *et al.*, 2011). Für die statistische Auswertung und Charakterisierung der untersuchten Gene wurden neben den Programmen „GENEIOUS“ und „MEGA5“ (Tamura *et al.*, 2011) auch die Programme der „EXPASY“ Internetseite genutzt. Die Identifizierung der Signalpeptide sowie ihrer Schnittstellen erfolgte unter Verwendung des „SIGNALP 3.0 SERVERS“. Für die Vorhersage der potentiellen N-Glykosylierungsstellen wurde der „NETNGLYC 1.0 SERVER“ der technischen Universität Dänemark verwendet.

## 2.7 Phylogenie

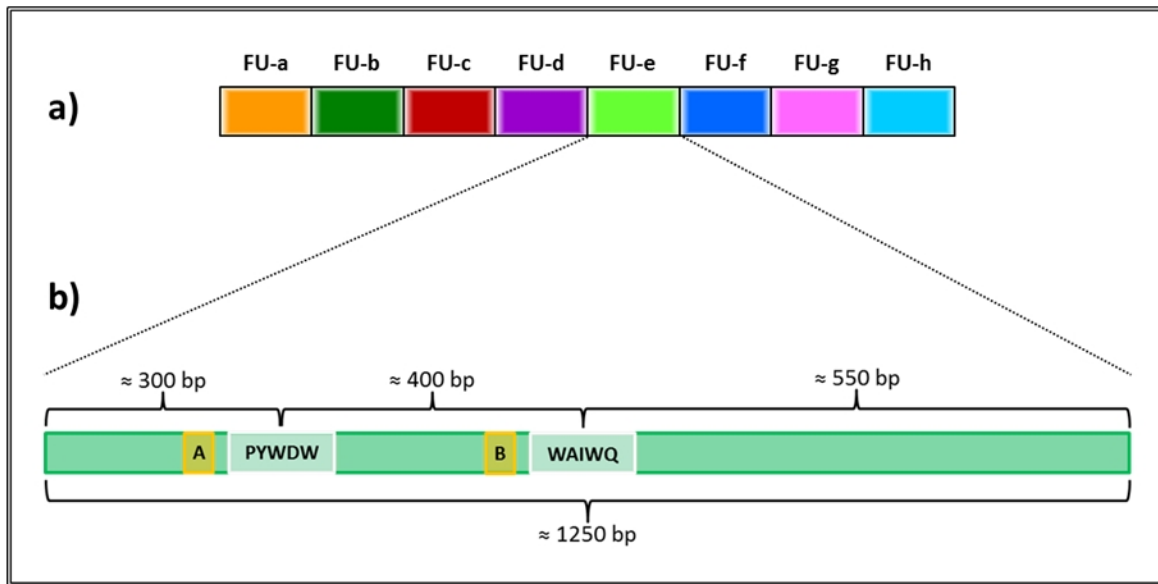
Für die Vergleiche von multiplen Aminosäure- oder Nukleotidsequenzen (Alignments) wurden die Programme „MAFFT“ (Kato *et al.*, 2002), „MUSCLE“ (Edgar, 2004), „MEGA5“, „CLUSTALX“, „GENEIOUS“ und die Internetressourcen „TRANSLATORX SERVER“ (Abascal *et al.*, 2010) sowie „PHYLOGENY.FR“ (Dereeper *et al.*, 2008) verwendet. Die graphische Bearbeitung der Alignments erfolgte entweder direkt in „GENEIOUS“ oder mit dem Programm „GENEDOC“. Alle durchgeführten phylogenetischen Berechnungen erfolgten unter Verwendung der Programme „BEAST 1.6.2“ (Drummond *et al.*, 2007), „GENEIOUS“ und „MEGA5“ sowie der Internetseite „PHYLOGENY.FR“. Für die graphische Aufbereitung und Auswertung der berechneten Stammbäume wurden die Programme „FIGTREE 1.3.1“, „TRACER 1.5“, „TREEANNOTATOR 1.6.2“ (Rambaut *et al.*, 2007), „TREEVIEW“, „MEGA5“, „GENEIOUS“ und „TREEILLUSTRATOR“ eingesetzt. Die für die einzelnen Berechnungen verwendeten Evolutionsmodelle und Parameter werden an den entsprechenden Stellen im Text aufgeführt und näher erläutert. Die Internetadressen aller in dieser Arbeit verwendeten Programme und Internetseiten sind im Anhang E aufgeführt.

## 3 Ergebnisse

Das Mega-Hämocyanin von *Melanoides tuberculata* (MtH) wurde erstmals in 2010 entdeckt und proteinbiochemisch charakterisiert (Lieb *et al.*, 2010). Diese Daten zeigten strukturelle Besonderheiten des Proteins, welche bisher noch bei keinem anderen untersuchten Hämocyanin beobachtet wurden. Neben den für Gastropoden-Hämocyanine typischen Dekameren wurden vorwiegend Tridekamere und Mega-Dekamere mit einem deutlich vergrößerten Kragenkomplex identifiziert. Auf proteinbiochemischer Ebene konnten zwei distinkte, am Aufbau des Mega-Hämocyanins beteiligte Untereinheiten unterschieden werden, die eine mit 400 kDa und die andere mit 550 kDa (Lieb *et al.*, 2010). Um die Struktur dieses neu entdeckten Hämocyanintyps aufzuklären, sollten die cDNA-Sequenzen, welche für die beiden Mega-Hämocyanin-Untereinheiten kodieren, im Rahmen meiner Arbeit ermittelt und analysiert werden.

### 3.1 Ermittlung erster Sequenzdaten

Durch frühere molekularbiologische Arbeiten am Hämocyanin von *Haliotis tuberculata* (Altenhein, 2000; Lieb *et al.*, 2000) konnten in unserer Arbeitsgruppe degenerierte Primer hergestellt werden, die es ermöglichten, spezifisch unbekannte Hämocyanin-Sequenzen zu ermitteln. Mit Hilfe dieser Primer, die innerhalb des Gens in der Nähe der hochkonservierten Kupferbindungsstellen Cu-A und Cu-B lokalisiert sind, konnten erste Sequenzdaten des MtH erhalten werden. Unter Verwendung der Primer PYWDW *forward* und WAIWQ *reverse* sowie genomischer DNA als Matrize konnten via PCR Fragmente mit der erwarteten Länge von etwa 400 bp generiert werden (siehe 2.5.7.2.).

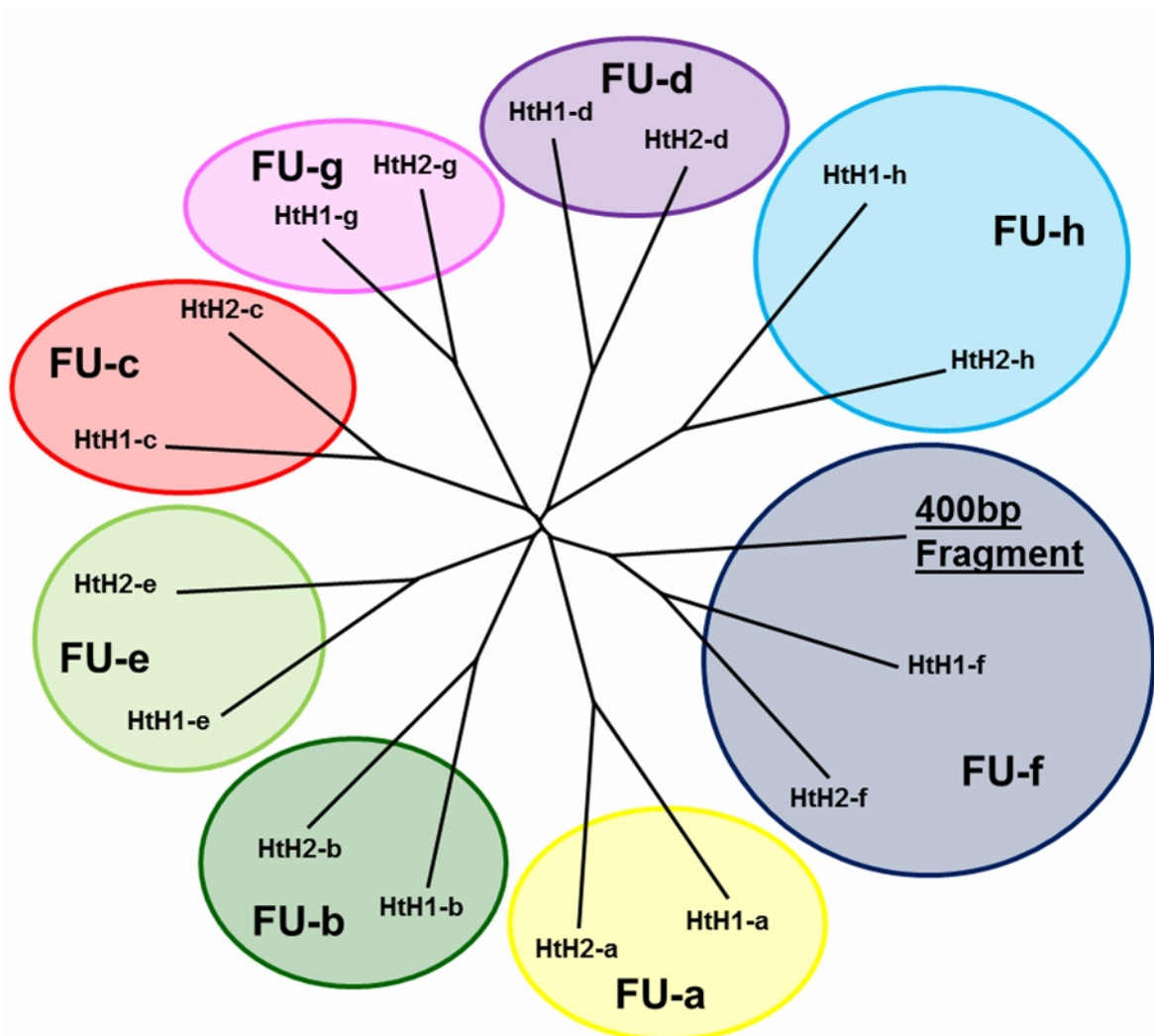


**Abb. 16:** Schema der Untereinheit eines Gastropoden-Hämocyanins sowie einer typischen FU

(a) Untereinheit eines Gastropoden-Hämocyanins bestehend aus den Domänen FU-a bis FU-h. (b) Schema der cDNA-Sequenz einer typischen FU. Die relative Lage der Bindungsstellen der degenerierten Hämocyanin-Primer PYWDW und WAIWQ (weiß) sowie die Lage der beiden Kupferbindungsstellen Cu-A und Cu-B (orange) sind durch Rechtecke gekennzeichnet.

Die PCR-Produkte wurden nach Gelelektrophorese und anschließender Gelextraktion (siehe 2.5.8) kloniert (siehe 2.5.9) und die erfolgreiche Klonierung mittels Klon-PCR verifiziert (siehe 2.5.7.6). Anschließend wurden 14 individuelle Klone sequenziert (StarSEQ<sup>®</sup>, Mainz) und die erhaltenen Sequenzdaten mit Hilfe des Programms „FINCHTV“ analysiert (siehe 2.6.2). Nach visueller Identifikation und Entfernen der Sequenzreste des verwendeten Plasmidvektors aus den Daten wurden die bereinigten Nukleotidsequenzen in ihre abgeleiteten Aminosäuresequenzen übersetzt. Um die so vorbereiteten Daten zu den acht funktionellen Einheiten eines typischen Gastropoden-Hämocyanins zuordnen zu können, wurden diese mit den bekannten Aminosäuresequenzen des Hämocyanins von *Haliotis tuberculata* (HtH1 und HtH2; Lieb *et al.*, 2000, 2001a; Altenhein *et al.*, 2002) verglichen. Dafür wurde mit Hilfe des Programms „MUSCLE“ (Edgar, 2004) ein multiples Sequenzalignment erstellt. Dieses wurde verwendet um einen einfachen radialen Stammbaum zu berechnen (Abb. 17). Da das typische Hämocyanin-Gen mit seinen acht FUs vor der Aufspaltung der Mollusken evolvierte, war es möglich, mit vorangehend beschriebener Methodik eine eindeutige Zuordnung der generierten Sequenzdaten zu den verschiedenen FUs zu erreichen. In den so berechneten Stammbäumen gruppieren die orthologen FUs der unterschiedlichen Hämocyanine immer zusammen in einem Ast (Lieb *et al.*, 2000; Lieb & Markl, 2004). Der in Abb. 17 dargestellte Stammbaum zeigt die Zuordnung eines mit den PYWDW- und WAIWQ-

Primern generierten 400 bp-Fragments zu den FUs von *Haliotis tuberculata*. Es ist klar zu erkennen, dass das entsprechende Fragment eindeutig der FU-f zugeordnet wird.

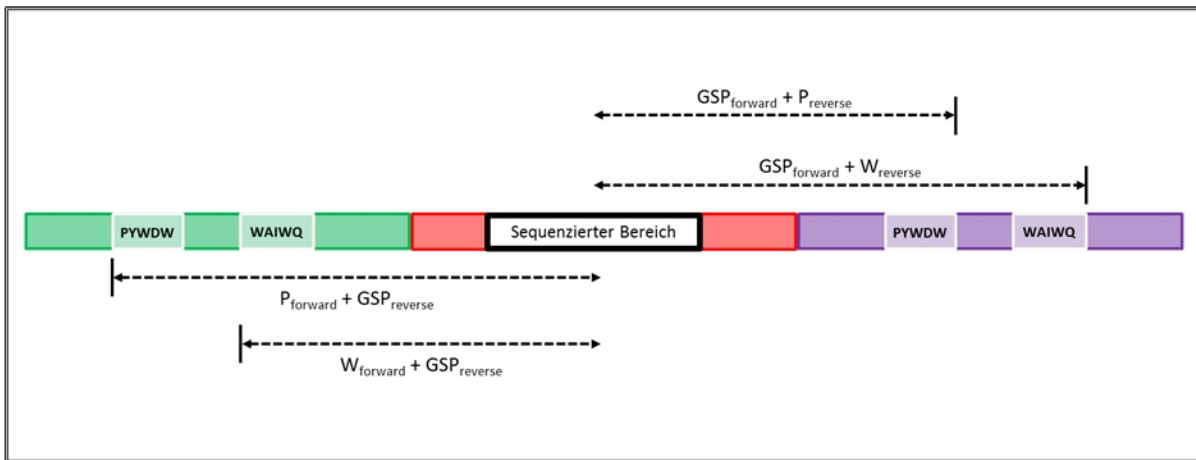


**Abb. 17:** Radialer Stammbaum der FUs der beiden Isoformen des Hämocyanins von *Haliotis tuberculata* (HtH1 und HtH2) und eines 400 bp umfassenden Sequenzabschnitts des *Melanoides tuberculata*-Hämocyanins (MtH)

Die Abbildung zeigt einen radialen Stammbaum basierend auf einem Aminosäurealignment (MUSCLE) der FU-a bis FU-h des HtH sowie eines 400 bp umfassenden Sequenzabschnitts des MtH. Die orthologen FUs des HtH gruppieren jeweils in einem Ast des Stammbaums. Deutlich erkennbar ist die Zuordnung des 400 bp Fragments des MtH zu den FU-f des HtH (FU-a = gelb, FU-b = dunkelgrün, FU c = rot, FU d = violett, FU-e = hellgrün, FU-f = dunkelblau, FU-g = rosa, FU-h = hellblau).

Die auf diese Weise erhaltenen DNA-Fragmente wurden anschließend genutzt, um genspezifische *forward*- und *reverse*-Primer zu generieren (siehe 2.5.7.1). Durch die Kombination der neu-synthetisierten Primer mit den bereits vorhandenen degenerierten Primern konnte das anfängliche 400 bp Fragment sowohl in 5'- als auch in 3'-Richtung verlängert werden (siehe Abb. 18). Aus den so erhaltenen DNA-Fragmenten konnten

wiederum neue genspezifische Primer hergestellt werden, die in weiteren PCRs erneut mit den degenerierten Primern kombiniert wurden.



**Abb. 18: Strategie zur Amplifikation der Hämocyanin-Gene von *Melanoides tuberculata***

Der initial sequenzierte Bereich des Hämocyanin-Gens von *Melanoides tuberculata* wurde verwendet um genspezifische Primer (GSP) herzustellen. Diese konnten in weiteren PCRs mit den bereits vorhandenen degenerierten Primern „PYWDW forward“, „PYWDW reverse“, „WAIWQ forward“ und „WAIWQ reverse“ kombiniert werden, um so das anfängliche 400 bp Fragment in 5′- und 3′-Richtung zu verlängern. „PYWDW“ und „WAIWQ“ kennzeichnen die Bindungsstellen der degenerierten Primer an diesen konservierten Sequenzmotiven.

Mit Hilfe dieser als „Primer walking“ bezeichneten Strategie war es möglich, große Teile der kodierenden DNA-Bereiche der Hämocyanin-Gene von *Melanoides tuberculata* zu amplifizieren und anschließend zu sequenzieren. Es wurde darauf geachtet, dass die 100 prozentigen Überlappungen zwischen den einzelnen DNA-Fragmenten eine Mindestlänge von ~100 bp hatten. So konnte eine sichere Zuordnung der Stücke zueinander garantiert werden. Die erhaltenen Sequenzdaten wurden anschließend mit Hilfe des Programms „GENEIOUS“ (Drummond *et al.*, 2011) zu größeren, als Contigs (von engl. contiguous = zusammenhängend) bezeichneten Sequenzen, assembliert.

## 3.2 Sequenzierung der Mth<sub>400</sub>-Untereinheit

Die nicht wie unter 3.1 beschriebenen amplifizierbaren 5′- und 3′-Bereiche der für die 400 kDa Untereinheit des Mega-Hämocyanins kodierenden cDNA wurden durch den Einsatz des „GeneRacer™ Kits“ (siehe 2.5.7.7) erarbeitet. Durch eine Vielzahl von PCR-Reaktionen, die unter Verwendung von 70 neu generierten, genspezifischen Primern (siehe Anhang F.4) durchgeführt wurden, konnten annähernd 300 einzelne Sequenzierungsansätze zur Bearbeitung an die Firma StarSEQ® übergeben werden. Der Großteil der

Sequenz der MtH<sub>400</sub>-cDNA war nach der Assemblierung der Rohdaten durch mehrere redundante Sequenzierungen abgedeckt.

```

MVTSPRFALALLAFCVQVAFTTGSLIRKKNVENLSSQEILSLQTSRLRALAADESASGFKAVAAAYHGEPSMCECDNDGVAIACCLHG
MAVFPQWHRLYVVQFEQLLVNKGKLNIGVYWDTELEPLTHPLSLVEEQIFRDPNGGKNSWFSGEVVPDGAIKNTARAV
DDRLFEEKVAPGKNTQLFNILDALEQDDYQCFEVQLEVAHNHIIHYLVGGRRHQYSMATLEYSSYDPIFFLHHSNIDRIFVIWQEL
QKKRGKPYDHSDCSVEQFRWNLKPFDRDTPNIALTKQFHTAKELFSLNLQGLVYVDDLTNLGMDVDQLYNLMEQRRARERA
FGVFSLHGLGFSANARVVKCSDLTGRRKRTIGSADLVEHHCEFAAGDFLLGGANEMPWEFHLPPYFDVSDAVQRLGLSLTSK
WSLEWDLYHVNGTALPGNLLPTPYGKHRPAKGRDSEEVSFADRDVHRHAVVRKNVVDLTDEEVYELRLAMDRLFQNDTS
VDGYQAIAEFHGLPARCPRDAAVRYACCIHGMATFPWHRLVQIEDELKARDLHFGMPYWDWTPNAQIPAIAAEAT
YQDPHHPDVTLHNPYDAVAFLNTRTRDVTQNDLAENPIYGDHSLYDGIILAFEQTDFCDFEIQFEVIHNSPHFLVGGFGA
YTLATLHYSADFPIFYLHHSNVDRMWAIWQELQIRRGPLPYRAHCAHTMTHEPMKPFSAAPLNNNEKTFHSEPSQVYDYE
RDLGYTYDSLTFGGMTIEQLENYLEERQSKERVVFGIELHNIGASAWVKVLIDSGSGGEPYLVGKIAVLGGEKEMPWHFDRLF
KMEITDALSDLGLSYDDNFVTLQITDVTGKEWSSDSFQHDITFHEPGKTFEVRGVSDKTRRLDVDHLNSEIQINLRDAFKAL
EDDTSINGFNQIAAFHGQPSWCPSDAETKYACCLHGMPTFPWHRLITVQLENALRDKGFEGLPYWDWTLPMTELPEF
TAETYDNPVSGANEANPLHHAHVVEGHKTRRSVREGELFQPPAFGKFRIGEQVMLAFEQTDYCSFEVQFEVAHNYIHALVGG
NELYSMASLRYTAYDPLFALHHSNTDRIWAIWQALQKYRGLPYNTANCAITQLRKLQPFQAQTSTVNPDSVTRDHSVPFDV
DYSSSFHYDYDNLEFNGLSIAQLQHEVVVRAKHERVFAGFMLHGKIQSALVRFEICSSGGACGQAGEFYILGDDYEMPWEYD
RLFKEYEITDQLEQGLEPLDRYDIKYTVFALDGSVVGDTDFGQVTVVHTYGAHVVERKSYGEEKASSHVRNVDDLGTGGEIES
LKAALQKMEDDGSFENIAEFHGAPGLCELNGKKGCCVHGMATFPWHRLYVEQVENALLSHGSAVSIPYWDWTKPITKLP
DIIAQETYFDSRSQTMQNNPFRGPIRTPDVNDYTRDPQPELFNNDYFLQQTLLALEQTSYCDFEAQFEVTHNAFHSFLGGR
GKYSLSTLDYSADFVFLYHANTDRIFAIWQALQSYRGLHWDEADCALNHMRDGLHPFDVKAQNEFDITHKYSRPVYDWP
YSDHLKYHDNLLNDWTIPQLEEVLTQQRSDRMFAGFLHNLIGTSADVELHICVATGNDRNCNPNPAGKFSILGGEYEM
FTDRLYKYDIDSDSIRKLGKLDSGANFDLKVQIHAYNGSYLDPSLLKEPTIIFEPGENTFQTQDGKTRQNMIRRNVLTLAERR
SLVLAMRRLQEDHSADGFQALASFHALPELCPYPEAAKRYACCVHGMASFPQWHRLYTVQFEEALRRHGSGLVGIPIYWDTSR
PQKALPAFATDEKFTDPVLNVEFDNPWLGADIEFENSHTEREPNARLGEEGEHGYDTWLYEQYLLALEQDNYCDFEVQFEIA
HNAIHAWVGGSKIHSMGLHYASYDPIFLHHSSTDRIFAVWQELQRLRGKDPNEANCALEIMHEPLKPFSGSPYNLNT
QRYSRPEDIFDYQSHFHYEYDSLEVLGMSPAQLEDAIHKRQEKDRVFAGFLKIGTSAHVEFTICNEAGVCTTAGDFDILGGS
AEMPWSFDRAKYDITNVLEKKGLEATDAFKVVAITAQDGTALDSSLPEPSIILQHATRDQGGYELAPNRVRRSLSDLTERD
VMSLSALHDLQEDDSATGWQSLAAYHGVPALCPSPEEAKYACCIHGMPTFPWHRLYVLAVEHSLIKHGSSVAIPYWDWT
NPLDKLPELFTKQTYDAWRDEVYDNPFGYVKGDEAYTVRDPQPELKERSRDGKHSVLFQVLLALEQTDYCDFEVQYEV
THNAIHYLVGGPQTFSLSSLHYSSYDPIFFVHHSFVDKIWAVWQALQRKRGKPSERADCAVNYMYDDMHFDSAKLDPDPT
IRSHANPSTVFNIIYDLGKYDDYQIGGKNLDELEALIKDNQAHPRVFAGFHLKIGSSADVVFICKICKDKSGGQCTRAGSIFILGG
EKEMPWEFDRLYKYDITDALKDNGIQPEDVFDAEAPFYLYEYVAVNGSTLPKSVINAPTLIFVPATGASVDAKSYAVAGMSV
RKDINTLSSAETQSLKNSLQKVMDEGEGPLSYQKIAEMHGWPGDCEQNGKVVACCHHGMAFPQWHRYVTRLLEMAMTW
QGANGVGIYWDWTEAFTALPSLVTDGDNPFNHGHIEALGKVTREPRPQLFKDPEHGEESFFYRQILLAFEQRDYCDFEVQ
FEVTHNAPHSWIGGTSYGMSTLEFSSYDPVFFLHHSNVDRQFAIWQALQKYRGLDYNSANCDIQELRMDLEPFNRDIPIVI
IRNNARAIDAFNYDQYQYDNLNFHGMISPELEALLEKRRSEHDVFLNFMLLGIGASADVTFDICSEGHCFAGTFAVLGG
PLEMPWSFDRLFYDVTSVFRQMHLRPSDYSVKVHVSNGTELDSDLLEAPSVSVFVPGKRGKAREDYRPTVSEDLRKEV
SSLSLEEVSNLKNALYKLNQNDHGPNGFEAIASFHGDPGLCEPGSGSKYACCQHGMPVFPWHRLTLVQFERALKAKGSSVGV
PYWDWTRPSRGIPQLFSNTYGNPNFLTYEIQAAAFVERDVQVDALEHTTGHEGFEDSLFHQALETLETSYCEFEVQFEMLH
NAVHALVGGTKTHSMATLEWSAFDPFFMVHHSIDRIWRIWQELQRLRKKPFNVARCAMRYLRKPLEPFSYASVNTDEVAR
TNSRPIDIFDTAKFHDFDNLDLGGHSIGEVNKMINDMRSKTRLVFVGLSGIETSATVKIEMDSSQGSHEVGTFFVLGGHN
EMPWAYERYVYKYEITAAKKYGLDHSVFNHFKVVKYDQGELSAQFKAPILVERPAGVDYDVAIFQLKKNLNDLAPKVVLP
GTRIRFSAEDGVSSVLREMGSYTNAVYCSIPPGDANTYKLEVEYTLPGDYFFVSNDEEKCKQGDRIQISIDEE

```

**Abb. 19: Aminosäuresequenz der MtH<sub>400</sub>-Untereinheit**

Die Abbildung zeigt die vollständige Aminosäuresequenz der 400 kDa Untereinheit des MtH. Die assemblierte cDNA wurde in die abgeleitete Aminosäuresequenz übersetzt und die einzelnen FUs farblich gekennzeichnet (FU-a = orange, FU-b = dunkelgrün, FU-c = rot, FU-d = violett, FU-e = hellgrün, FU-f = dunkelblau, FU-g = rosa, FU-h = hellblau). Das 23 Aminosäuren umfassende Signalpeptid am 5'-Ende ist schwarz dargestellt. Die gesamte Sequenz umfasst 3445 Aminosäuren. Die gesamte Untereinheit besitzt eine theoretische Molekularmasse von 390 kDa.

Um die zusammengesetzte Sequenz abschließend zu verifizieren, wurde diese in acht überlappenden Fragmenten von jeweils etwa 2000 bp Länge amplifiziert und sequenziert.

Die Überlappungen waren mit etwa 1000 bp ausreichend groß bemessen, um eine sichere Assemblierung der cDNA zu gewährleisten.

Abb. 19 zeigt die so ermittelte vollständige Mth<sub>400</sub>-Untereinheit nach Übersetzung der cDNA in die abgeleitete Aminosäuresequenz. Sie besteht aus insgesamt 3445 Aminosäuren und beinhaltet am 5'-Ende ein Signalpeptid von 23 Aminosäuren Länge. Die berechnete Molekularmasse der Mth<sub>400</sub>-Untereinheit beträgt 390 kDa. Ein Schema der Mth<sub>400</sub>-cDNA ist in Abb. 20 dargestellt. Sie umfasst insgesamt 10915 bp. Zu Beginn der Sequenz liegt eine 5'-UTR von 99 bp. Daran anschließend folgt der offene Leserahmen (open reading frame = ORF) mit 10335 bp, der das Signalpeptid (69 bp) und die acht paralogen funktionellen Einheiten FU-a bis FU-h (10266 bp) umfasst, die typischerweise bei Gastropoden-Hämocyaninen vorzufinden sind (Lieb *et al.*, 1999). Abschließend folgt die 3'-UTR mit einer Länge von 481 bp sowie der typische poly(A)-Schwanz. Die vollständige Nukleotidsequenz der Mth<sub>400</sub>-cDNA ist in Anhang dieser Arbeit (G.1) abgedruckt.



**Abb. 20: Schema der Mth<sub>400</sub>-cDNA**

Dargestellt sind die acht konkatenierten funktionellen Einheiten (FU-a bis FU-h) der für die 400 kDa Untereinheit des Mth kodierenden cDNA. Die Abkürzung „SP“ bezeichnet das Signalpeptid am 5'-Ende. Die untranslatierten Bereiche (untranslated region = UTR) an beiden Enden der cDNA sind mit schwarzen Quadraten kenntlich gemacht.

### 3.2.1 Die 5'-UTR der Mth<sub>400</sub>-cDNA

Upstream des translatierten Bereiches der cDNA schließt sich eine kurze 5'-UTR von 99 bp Länge an. Der GC-Gehalt der 5'-UTR ist mit 52,5 % nur geringfügig höher als der AT-Gehalt. Die häufigsten Nukleobasen sind Cytosin und Thymin (Adenin: 14,1 %, Cytosin: 31,3 %, Guanin: 21,2 %, Thymin: 33,3 %).

### 3.2.2 Die 3'-UTR der Mth<sub>400</sub>-cDNA

Am 3'-Ende schließt sich an die kodierende Sequenz der Mth<sub>400</sub>-cDNA ein 481 bp umfassender untranslatierter Bereich (UTR) an, der mit dem Stop-Codon (TAG) beginnt. Daran angrenzend folgt der charakteristische poly(A)-Schwanz. Ein Polyadenylierungssignal (5'-AAUAAA-3') findet sich 22 bp upstream des poly(A)-Schwanzes. Zusätzlich sind weitere Motive der Form „UGUAN“ vorhanden, welche die Bindung des „cleavage

factor I“ (CFI) an die RNA im Rahmen der Polyadenylierung ermöglichen (Venkataraman *et al.*, 2005). Die Basenzusammensetzung und damit auch der GC-Gehalt der UTR ist ausgeglichen (Adenin: 25,6 %, Cytosin: 22,5 %, Guanin: 25,6 %, Thymin: 26,4 %; GC-Gehalt: 48 %). Die 3'-UTR der Mth400-cDNA ist in Abb. 21 dargestellt.

```

5'-TAGATGGCGGATGCTGTATCAGACAACAGACTTTATTTCATTCTTGGAGACTTTACAAACC
ATGTTATAGCCTTGACTTGCCTTGCCTGTCACTGTCACTTGTGCATCTACTCCCAGAGTTTCAAAGAT
TTAACCTTTTACCTCCAGCCGCGGTTATGGCCCATGCCGAACGCATAGGCCATAACCCCGG
CTGGAAGTGAAAGATTTATCTCTTGACCTCCTCGAAACTGTACTCCGCCATGCTGGAAGCTT
GGTAGCTACAGATTCTGGGAGCACACATATTTCACTACTATAGTCACCTGTGGCCCATGCTG
AAAGCATGGGCCACAGGCGACTATGGTGGTGAAGGGTGAAGAGGTAGCTTCGCCACATC
ATAGGGTGGAGATACCGAGTCTTTGAAGCTGCTGGAGTAGCAGTGTGGATGGGAGATTTTA
GATAGTGTGTGAGTGTGGTGGAAACTGCATAAACTCAGTGCAATACTGTAAAAAAAAA
AAAAAAAA-3'

```

**Abb. 21: 3'-UTR der Mth<sub>400</sub>-cDNA von *Melanoides tuberculata***

Die Abbildung zeigt die vollständige Sequenz der 3'-UTR der Mth<sub>400</sub>-cDNA. Sie beginnt mit dem Stop-Codon „TAG“ (rot) und endet mit dem poly(A)-Schwanz (orange). Das Polyadenylierungssignal „AATAAA“ ist in blau dargestellt und befindet sich 22 bp upstream des poly(A)-Schwanzes. Die vorhandenen „UGUAN“-Motive (Venkataraman *et al.*, 2005) sind in einer grünen Farbe eingezeichnet.

### 3.2.3 Die Mth<sub>400</sub>-cDNA besitzt ein Signalpeptid

Direkt zu Beginn des translatierten Bereiches der Mth<sub>400</sub>-cDNA befindet sich ein für Mollusken-Hämocyanine charakteristisches Signalpeptid von 23 Aminosäuren Länge (siehe Lieb *et al.*, 2001). Die Identifikation der Schnittstelle, an der die Leader-Peptidase (Dalbey *et al.*, 1997) das Signalpeptid abspaltet, erfolgte mit Hilfe des „SIGNALP 3.0 SERVERS“ (Bendtsen *et al.*, 2004). Die graphische Ausgabe des verwendeten Programms ist in Abb. 22 dargestellt. Die Schnittstelle befindet sich mit der höchsten Wahrscheinlichkeit an der Position, an der der C-Score („cleavage site“ score) den signifikant höchsten Wert erreicht. Ein hoher S-Score kennzeichnet diejenigen Aminosäuren, die mit hoher Wahrscheinlichkeit Teil des Signalpeptides sind. Das Signalpeptid der Mth<sub>400</sub>-Untereinheit beginnt mit der Aminosäure Methionin und endet mit den drei Aminosäuren Threonin, Threonin und Glycin. Die FU-a beginnt demzufolge mit der Abfolge Serin, Leucin und Isoleucin. Die Schnittstelle für die Signalpeptidase liegt zwischen den Aminosäurepositionen 23 und 24 und somit an der Stelle, wo der Abfall des S-Scores steil und ein signifikanter C-Score detektierbar ist. Zusätzlich zeigt der Y-Score, welcher ein Derivat des C-Scores in Kombination des S-Scores ist, durch einen deutlichen Peak die Schnittstelle an.

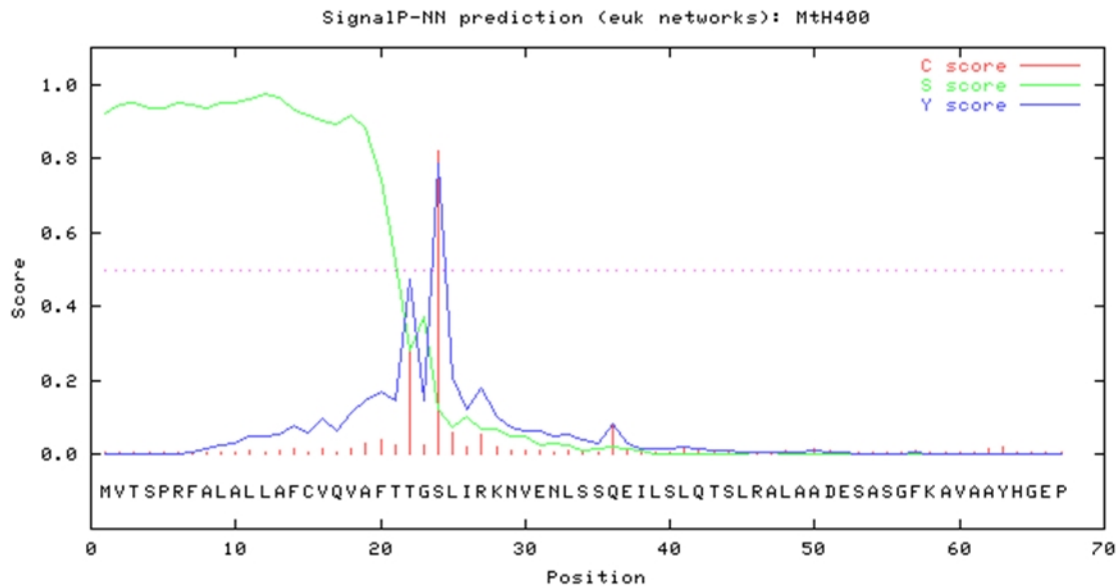


Abb. 22: Graphische Ausgabe des „SIGNALP 3.0 SERVERS“ für die Mth<sub>400</sub>-Untereinheit.

Die Abbildung zeigt im unteren Bereich die Aminosäuresequenz der Mth<sub>400</sub>-Untereinheit beginnend mit dem Start-Methionin. Die X-Achse zeigt die Positionen der Sequenz an, während die Y-Achse die Höhe der einzelnen Scores angibt. Die Schnittstelle für die Signalpeptidase findet sich an der Stelle, an der der S-Score steil abfällt und der C-Score („cleavage-site-score“) signifikant hohe Werte erreicht. Zudem ist der Y-Score an der Schnittstelle, welche hier zwischen den Aminosäurepositionen 23 und 24 (TTG-SL) liegt, deutlich erhöht.

### 3.2.4 Die Mth<sub>400</sub>-cDNA besitzt eine „Translation initiation site“

Die „Translation initiation site“ (TIS), die auch das Startcodon (AUG) miteinbezieht, folgt im Allgemeinen bei eukaryotischen mRNAs dem Schema „PuXXAUGPu“ („Pu“ = Purin, „X“ = beliebige Base, „AUG“ = Startcodon). Dabei ist als „Pu“ (Purin) in den meisten Fällen Adenin vorhanden (Szczesna-Skorupa *et al.*, 1986). Außer dem Purin auf Position -3 wird auch ein Purin auf Position +4 (vorzugsweise Guanin) mit einer höheren Translationseffizienz in Verbindung gebracht (Kozak, 1987). Eine typische Mollusken-TIS ist nach folgendem Schema aufgebaut (Mankad *et al.*, 1998):

| -6  | -5  | -4  | -3 | -2 | -1  | START |   |   | +4 |
|-----|-----|-----|----|----|-----|-------|---|---|----|
| a/t | c/t | A/C | A  | a  | a/c | A     | T | G | g  |

Abb. 23: Consensus TIS der Mollusken

Die Abbildung zeigt die Consensus TIS der Mollusken (Mankad *et al.*, 1998) von Position -6 bis +4. Dabei wird das bevorzugte Nukleotid an jeder Position angezeigt. Nukleotide in Großbuchstaben entsprechen der 50/57 Regel (Cavener, 1987). Kleinbuchstaben bezeichnen Nukleotide mit der höchsten Frequenz ohne jedoch der 50/57 Regel zu entsprechen. Erreichen zwei Nukleotide gemeinsam eine Frequenz über 75 % gelten sie als „Co-Consensus“ und werden gemeinsam erwähnt.

Die TIS der MtH<sub>400</sub>-cDNA weist die in Abb. 24 gezeigte Nukleotidsequenz auf.

| -6 | -5 | -4 | -3 | -2 | -1 | START |   |   | +4 |
|----|----|----|----|----|----|-------|---|---|----|
| C  | A  | C  | G  | T  | C  | A     | T | G | G  |

**Abb. 24: TIS der MtH<sub>400</sub>-cDNA**

Die Abbildung zeigt die TIS der für die MtH<sub>400</sub>-Untereinheit kodierenden cDNA von Position -6 bis +4. Das Startcodon „ATG“ ist rot hinterlegt. Die als wichtig für eine effektive Translation geltenden Purine an den Positionen -3 und +4 sind orange eingefärbt (Kozak, 1987).

### 3.3 Sequenzierung der MtH<sub>550</sub>-Untereinheit

Durch die proteinbiochemischen Untersuchungen von Lieb *et al.* (2010) war bekannt, dass das zentrale Mega-Dekamer aus einer zweiten, etwa 550 kDa großen Polypeptidkette aufgebaut ist (siehe 1.3). Es handelt sich hierbei um die zweitgrößte jemals bei einem Protein detektierte Polypeptidkette. Auch bei der Amplifikation dieser cDNA wurden die initialen Fragmente durch die Verwendung der degenerierten Hämocyanin-Primer gewonnen (siehe auch 3.1). Das weitere Vorgehen entsprach der Strategie, die auch bei der Amplifikation der MtH<sub>400</sub>-cDNA erfolgreich angewendet werden konnte (siehe: Abb. 18). Die besondere Schwierigkeit lag hier jedoch bei der Auswahl der zu verwendenden Primer. Durch die hohe Ähnlichkeit der beiden cDNAs galt es, Kreuzhybridisierungen der Primer zu vermeiden und entsprechende Oligonukleotide mit hoher Spezifität für den gewünschten DNA-Bereich zu finden. Dafür mussten teilweise mehrere verschiedene Primer für einen bestimmten Bereich generiert und getestet werden. Für die vollständige Sequenzierung der MtH<sub>550</sub>-cDNA wurden PCR-Reaktionen unter Verwendung von 87 verschiedenen Primern durchgeführt und abschließend die gesamte Sequenz aus annähernd 400 teils 40-fach redundanten Sequenzabschnitten assembliert.

MTWLAFALLVLGALVLSQSAALLVRKNVEDLTQEEVLSLQSYLSKLEADESEYGFQKLAAYHGPESMCQSDDGATARACCIHGMANFPQWHRLYV  
 QFEQLLRKKGMSMGVYWDVTGELTSPLDVEEHIFRDPNGGRGKENAWFSANIKHGNVDTKTARAVDKRFLQKVGDEKTLQFNILDALEKDD  
 YCQFEVQFEVAHNNHILVGGRRHKYSMSTLEYSYDPIFFLHHSAMDRIWVWQELQKRGKSSDHADCSVELFRWNLWPFDRSSNPVELTKQHSKA  
 DSVVHHPELGYWYDRTLGGMDLDQLHAYLEKRRARERAYATFSLHGVGYSANARVQVCGLSSEHPRDTRRRRSDASDNCFAGDFLLGGANEM  
 PWEFHLPYFDVTEAVQRLGVALDGAWTVKVDYHVNGTAMPGDILPPPVAHRPAKGAHDSDDIHGHSKSPDIIVRKDVSELTAEMEYELRQAM  
 SKFQNDTSVDGYQVAEAFHGLPARCPRPDAAVRYACCIHGMAFPWHRLVQIEDELRLARGLHIGVYWDWTRPNTHIPKLAEEATYQDPHHE  
 GVTLNPNFYEAFAFLGKTERDVQSQSENPPFGDHTDLDFGMLLAFEQDFDFEIQFEVTHNAPHFLVGGFKPYSLATLHYSADFIFYLHHSNV  
 DRLWAIWQELQRRGLPYEAHCANSMTDHLKPFGEAPLNNNEKTHSHARPTKIYDYELELDYTDNLQFGGMTIEQLEHYLEERQARERFVGLF  
 FHNIGVSAWADVYVAGGQRYRAGGVAVLGGEQEMPWFHDFRLFKMEITDALKALGLGYNDFTISLNTDVGNGKHVDENTFSHDTIFHDPHVVH  
 HDAGDAPSPDKVRSNIESLREQVQNLRDALQALEDDSSNAGFAQIAAFHGEPRWCPSTDAEKKYACCLHGMAVFPWHRLITVQAENALRHGF  
 EDGLPYWDWTRPMKALPQVVEAQQYVYVQQTGQTMNDNPFHGHGIANHATERDVSPKLFQPEFGKFTDIAELVMFAFEQDFDFEIQMEIAHN  
 YIHALVGGSKPYSMSSLRYTYDPIFFLHHSNTDRIWAIWQALQKRYRGLPYNTANCALASMRKPLQPFQAQTSVVPNDPMTDRHVSVPFEIDYQSSFH  
 YTYDDLEFNGLSIPQLQHEVVVQAHERVFAFGFMLKIGIGQSAVLTVEICQSSGQCSAAGEFYMLGDEYEMPVWVYDRLFKYEITQLQADKGLKPLDSY  
 YVQYQVKDLQNNVIGTDNFGKVTIVHTYGAHGLDRKSYGQELKASHVRQNVDSLSEGGKESLKAALQKMSDGSFEKIAEFHGAPGLCKHEGRNV  
 GCCHIGMATFPWHRLYVEQVENALHTHGEVSIYWDWTSPIKELPDLVSKATYFDSRSQTMETNPFKAQIGNTGHYITRDRPQPLFNTKYFLN  
 NVLLALEQTSFCDEAQFEVHNMMHSFVGGRGKYSLSTLDYSADPLFFIYHANTDRIWAIWQALQSRGLPWDESDCALNHMHDLHPFDDK  
 AQNEFDITHKYNHPDDVWYSDHLDYHYDTLELNGWSIAELEVELEKQSRDRMFAGFMLHDIGTSADVEIFICLAAGKNQRDCNHPAGKFSVLGG  
 EFEMPFADFRQYKFDISEALTKLIRGLSNADFDLKVTVQSYNGSYLSSDLDAPTIIFSPGNRKVQSDRPAAADGDLVRRNVWLSLQFERRSLVLAIR  
 LQEDHSADGFQSLAFHAEVPELCPYEPAPKRFACCVHGSFAPFHWHRLYTVQFEALKRHGSLVGLPYWDTTDAQSFIPNFLTDEKFDHTVDDVDI  
 PNPWLGADIEFENSHTERQFNMDKLRKEGPKGYDTSWSEQVMLAFEQENCFDFEVQFEVMHNAIHAWTGGTSVHSMALHLYASYDPIFMLHHS  
 STRIFAIWQELQRLGLDPSANCDLETMHAPLKPFSFGSPYNLNPDKRYSSPDIFDYHGHFHYEYDDLEMQLSPPQILDRIAQRKERERVFAG  
 FLLHGIGTSAHVFSVCKDNACTPAGDFNLGGYAEEMPWSFDRAYKYEITDVLKKNIDPRDQFTIELKITALNGTTLISGNAFGEPSIIQARS  
 AERGGY  
 NQPANRIRHDLTHLSEDFMSLSALNDLQADTGKMGWQSLAFHGVPAALCPTAEATYACCIHGMAFPWHRLYTLLEHALIEHGSSAIYVW  
 DWTQPLDGLPALFTEQTYDAWKDKVDFNPFARGYKISVDGYTVRDPQSQMSMRKDKGHSILFDRVLDALQDFDFEIQFEVTHNAIHYLVG  
 GPQTYLSLHYSSYDPIFFVHHSFVDKIWAVWQELQRRRGKPYDRADCAVNYMSKPMMPFNSPDLDPDLEVRHAHVNSVFNELGYKYNLEI  
 GGESLEQLEDRIQQSHARVFAFYLGIGTSAGVTFVAVCRKDDGKRCYTGSSFFILGGSKEMPWSFDRVYKHDMTALHAIGLEPEDVFDYKYN  
 VFISYNTAVNNTALPSSIIKPTILYEPKVPANRVRKDLTDLSEADVNSLSALAEVQADRGLGWQSLASYHGVPAALCPTPEAAAYACCVHGMPT  
 FPHWHRLYTLVDRALIHGSRVAIPYWDWTLPIKELPAITAKTFHDPDADVPNPFARAYIAEANTYIRRDQPPELKRVTKDGEHNVLFDEVLAL  
 EQDDFCDFEIQFEVTHNAVHYLVGGHDEHSLASLHYSSYDPAFFLHHSFVDKIWAVWEEELQKRRHKPYVSADCAAKYLNELPHFPDDPDLDPDDY  
 VRTHAMPSSDSFDYHNLGYDYYDDMKIGGKSLDQLADIIKEHQSHSRAGFHLEGIGASADVEFEVCRGGQCQKAGTFFLLGGAKEMPWSFDRLYK  
 DITDPLHALGIEPEDAILSPAPFTFKIQVFALNGSSLPSSTLDAPKVLVYAPAAANRMRSSVDLSNAEATTSREAFKMRDYDTPGLGSQALASYHGLPAM  
 CPTAEATYACCIHGMAFPWHRLYTLLEVDRLVSHGSSVAVPYWDWTLVKKLPALIDDATYFNDSNGKQENPFNNGYIMEAKTKTVRDPQ  
 ELFKVSKDQHSVLFDEVLLEQDDYCDFEIQFEVTHNALHYLIGGRNLYLSSLHYSSYDPAFFMHHSFIDKIWVWQKQKRRGKEYQAKCAK  
 YMSEALHPFDDPELDPDELVRTHAMPDDVFNMMALGYKYNLEIGGKNLDELEAIKHEQSHSRVFAFGHLEGFQTSADVVINCRDDTCARAGAF  
 FVLGGSKEMPWSFDRQFKYEITDALAELNIEPADVTDVCRPPFSKLDIHAINGSALSSVLDAPTIVVQSAVYVDERVRHDLNHMSAKDLVSLRNLAL  
 ELNDDTGLVGFQSLASFHGVPAALCPTPEAAAYACCIHGMAFPWHRLYTLLEVERALUWHGSKVSMYVWDWTLVSELPLATEQSYDYPKSETR  
 YENPFSRAYIGEARAYTVRDRPELTKVTKNGANHVLFDEVMLEQEDYCDFEIQFEVTHNAMHYVGGHQLHSLASLHYSSYDPIFFLHHSFVDKI  
 WVVWQELQKRRHKPYNTAECAAKYLAEPHFPDDPDLDPDDVVRTHAMPNDVFDINELGYKYNDFEIGGKDLQLEAIKQRRSHDRFAGWRK  
 GIKTSADVLFNVCTKNQGSSENCVRGGKILLGGEKEMPWAFDRPYLVEITELQGLGIPVEEASKTSFYSFYFTWEVIGDGHSLPASSLSPPTAVFKAG  
 YDFPSERVHESDLSESRLSKTALVDLQDEGKLGWQSLAAYHGVPAALCPTPEAATYACCVHGMPTFPWHRLYTLLEVDRLALRHGSKVAIPYW  
 DWTLVPTLPSILTSKTFNDDHGRMESNPFASKYISESSYIQRDPQPELFFKSKDKGHSVLFDEVLVYALEQQDYCDFEIQFEVTHNAFYMLGGFNE  
 HSLSSLHYSSYDPAFFLHHSFVDKIWAVWQKQKRRGRPYNTAQCAADYMDPELHPFDDPVLDPDELVRSHALPTDVFYDHLGYTDDLSIGGLSP  
 SQLADLIAKEGHDRVFAFGFLAGIGTSKAVFDFVCAGAGDSNCRRAAGFLLGGSEMPWSFDRPYRYEITDILDELHIEPEDATDKVPPFHRFRIT  
 ALNGSALSSDILGTPKIIFVGVHVPDRVRHDLRSVSEDLRSLKLSALGELLEDDGKMGWQALASYHGVPAALCPTPEAATYACCIHGMAFPWHRL  
 YTLVVDHALYRHGSKLSPYWDWTLPMDKLPSILTADTYNENHAKREANPFLGAYINESHQYIVRDPQPALWKVSGDGEHSLVFEVLYALEQDDF  
 CDFEVQFEVTHNALHYLVGGHRLYLSLSSLHYSSYDPAFFLHHSFVDKIWAVWQELQKRRGKAYNRADCAARFMGEPLHPFDDPDLDPDEIIRTHAM  
 PNDVFDIHELGYSDNLDIGGHSIEELERHIANHQNHDRVFAFGRLSGIKTSADVEFFVCKGPDTTGAKSDENCVRGGGFFLLGGEKEMPWSFDRLYK  
 YEITDQVRELGINPSDAAGEDPGFHLHVTAHADINGDEVPASALAKPEVFDAAHLAPVPRDRVRYELSSVSEDLSSLKAALGELQHDGKLGWQSL  
 ASYHGVPAALCPTPEAATYACCVHGMPTFPWHRLYTLLEVDHALFRHGAKVSIYWDWTRQMDRLPSILTRASFFNEESGQSEPNPFLKAYINESHQ  
 YTVRTPQSAWVWSSDQHSVLFDEVLVYALEQDDFCDFEIQFEVTHNAIHYLVGGHRLYLSLSSLHYSSYDPAFFLHHSFVDKIWAVWQELQRRRGK  
 YNTADCAKYLNEPLHPFDDPDLDPDEDVRAHAMPSDVFYRGEGLYDYNFNIGGKSINQLEALIRQRHGHDRAFAEFSLGQKTSADVEFFVCKN  
 DNCVRGGGFFLLGGEKEMPWSFDRPYKYEITDQLKELGIEANEATIRVNVHDVSGAELASVGTPRITVDPAH

Abb. 25: Aminosäuresequenz der Mth<sub>550</sub>-Untereinheit

Die Abbildung zeigt die vollständige Aminosäuresequenz der 550 kDa Untereinheit des Mth. Die assemblierte cDNA wurde in Aminosäuren übersetzt und die einzelnen FUs farblich kenntlich gemacht (FU-a = orange, FU-b = dunkelgrün, FU-c = rot, FU-d = violett, FU-e = hellgrün, FU-f = dunkelblau). Die insgesamt sieben FU-f wurden jeweils abwechselnd rot unterstrichen um somit eine einfachere visuelle Abgrenzung zu ermöglichen (rot unterstrichen: FU-f<sub>1</sub>, FU-f<sub>3</sub>, FU-f<sub>5</sub>; ohne Unterstrich: FU-f, FU-f<sub>2</sub>, FU-f<sub>4</sub>, FU-f<sub>6</sub>). Das 20 Aminosäuren umfassende Signalpeptid am 5'-Ende ist schwarz dargestellt. Die gesamte Sequenz umfasst 4999 Aminosäuren. Die gesamte Untereinheit besitzt eine theoretische Molekularmasse von 567 kDa.

Die außergewöhnlich hohe Redundanz beruht unter anderem auf falschen Primerbindungen, die somit zur mehrfachen Amplifikation ein und desselben Bereichs führten. Auch bei der Mth<sub>550</sub>-cDNA erfolgte die Amplifikation der 5'- und 3'-Bereiche unter Verwendung des „GeneRacer™ Kits“ (siehe: 2.5.7.7). Die korrekte Assemblierung der Mth<sub>550</sub>-cDNA wurde abschließend durch die erneute Amplifikation der gesamten Sequenz in acht Fragmenten mit jeweils 2500 bis 4000 bp Länge verifiziert. Es wurde darauf geachtet, dass die Überlappungen zwischen den einzelnen PCR-Produkten etwa 1000 bp betragen und damit eine sichere Zuordnung der Sequenzen zueinander ermöglicht wurde. Die abgeleitete Aminosäuresequenz der vollständigen Mth<sub>550</sub>-Untereinheit ist in Abb. 25 gezeigt. Sie besteht aus insgesamt 4999 Aminosäuren und beinhaltet am 5'-Ende ein Signalpeptid von 20 Aminosäuren Länge. Die rechnerisch ermittelte theoretische Molekular-masse der Mth<sub>550</sub>-Untereinheit beträgt 567 kDa. Ein Schema der entsprechenden cDNA ist in Abb. 26 dargestellt. Die Sequenz umfasst insgesamt 15346 bp und beginnt am 5'-Ende mit einer UTR von 75 bp Länge. Daran anschließend folgt der offene Leserahmen (14997 bp), der aus dem Signalpeptid (60 bp) und der für die Mth<sub>550</sub>-Untereinheit kodierenden Sequenz (14937 bp) besteht. Abschließend folgt die 3'-UTR mit einer Länge von 274 bp sowie der typische poly(A)-Schwanz. Die vollständige cDNA-Sequenz der Mth<sub>550</sub>-Untereinheit ist dem Anhang der vorliegenden Arbeit zu entnehmen (G.2).



**Abb. 26: Schema der Mth<sub>550</sub>-cDNA**

Dargestellt sind die zwölf konkatenierten funktionellen Einheiten (FU-a bis FU-f<sub>6</sub>) der für die 550 kDa Untereinheit des Mth kodierenden cDNA. Die Abkürzung „SP“ bezeichnet das Signalpeptid am 5'-Ende. Die untranslatierten Bereiche (untranslated region = UTR) an beiden Enden der cDNA sind mit schwarzen Quadraten kenntlich gemacht.

### 3.3.1 Die 5'-UTR der Mth<sub>550</sub>-cDNA

Am 5'-Ende der Mth<sub>550</sub>-cDNA befindet sich eine UTR von 75 bp Länge. Der GC-Gehalt ist mit 61,3 % deutlich höher als der AT-Gehalt. Dementsprechend liegt die prozentuale Häufigkeit der Nukleobasen Cytosin (36 %) und Guanin (25,3 %) in der cDNA über der Häufigkeit von Adenin (17,3 %) und Thymin (21,3 %).

### 3.3.2 Die 3'-UTR der Mth<sub>550</sub>-cDNA

Der ORF der Mth<sub>550</sub>-cDNA endet vor dem Stop-Codon „TAG“. Danach folgt die 274 bp umfassende 3'-UTR. Der GC-Gehalt der UTR ist mit 45,6 % leicht erniedrigt was sich auch in der Häufigkeit der vorkommenden Nukleobasen widerspiegelt (Adenin: 28,1 %, Cytosin: 19,7 %, Guanin: 25,9 %, Thymin: 26,3 %). Ein Polyadenylierungssignal, welches dem typischen Hexanukleotid „AAUAAA“ entspricht (Wahle *et al.*, 1992) ist in der Mth<sub>550</sub>-cDNA nicht identifizierbar. Stattdessen findet sich eine Variante des erwähnten Motivs, nämlich „UAUAAA“, in einer Entfernung von 69 bp upstream des poly(A)-Schwanzes. Das „UAUAAA“-Motiv wurde bereits in einer groß angelegten Studie als dritt häufigstes Polyadenylierungssignal bei Genen von Mensch und Maus identifiziert (Tian *et al.*, 2005). Zusätzlich finden sich weitere „UGUAN“-Motive in der 3'-UTR der Mth<sub>550</sub>-cDNA. Bei Vertebraten konnte gezeigt werden, dass die „UGUAN“-Motive selbst bei fehlenden „AAUAAA“-Signalen eine effektive Polyadenylierung ermöglichen (Venkataraman *et al.*, 2005). Die Position der Signale in der 3'-UTR ist in Abb. 27 dargestellt.

```

5'-TAGTGTAAAAAAAAAAAGTAGTAACCACTAGACAACCTACGCGTGTAGTCAGGTCATCA
TTGTATGTTGTAACAGCTCATTCCATTGGACGGCCTGCTTGCAAAGTGTGCATTGACAAGAC
TGATTGGTTGATCCGTTGAGGCCTGAAAGTATATGCTGGCCTTGCTGTTAGCATGGCCTGCA
GAGACTTGCAAGCCCCAAGAGATATAAAGGACTGAGGTTAGGTAGGGGAAAGGCTTGCAA
TTGACAGCCCCATTGTCACTTTGGTCTAGGTAAGGACTGAGGTTAGGTAGGGGAAAGGCTTGCAA
AAAAAAAAAAAAAAAAAAAAA-3'

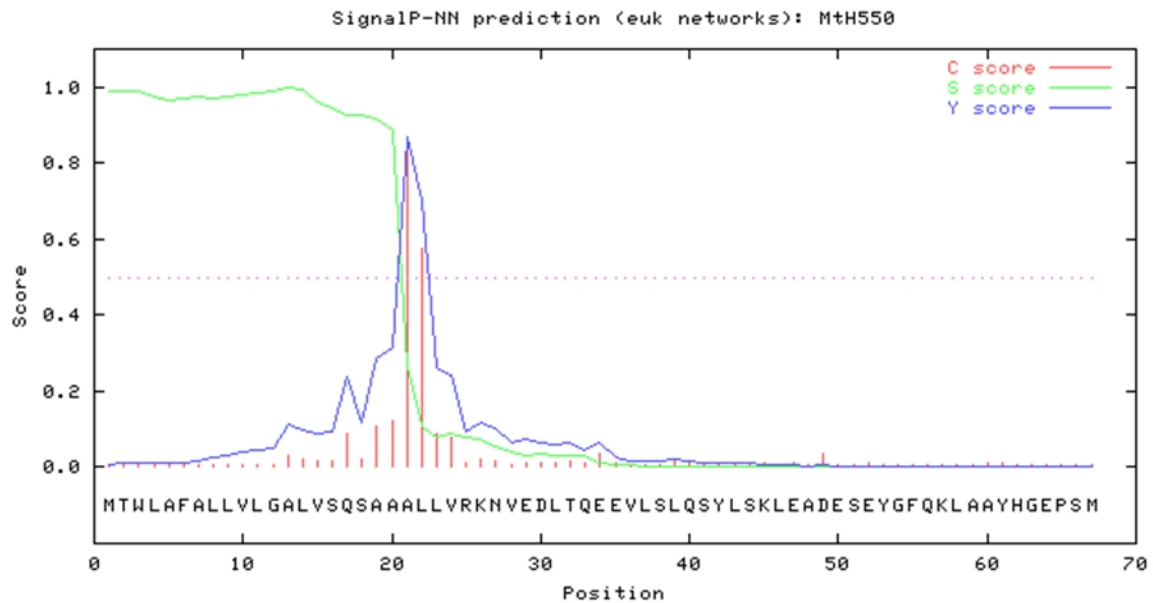
```

Abb. 27: 3'-UTR der Mth<sub>550</sub>-cDNA von *Melanoides tuberculata*

Die Abbildung zeigt die vollständige Sequenz der 3'-UTR der Mth<sub>550</sub>-cDNA. Sie beginnt mit dem Stop-Codon „TAG“ (rot) und endet mit dem poly(A)-Schwanz. Das durch einen Nukleotid-austausch veränderte Polyadenylierungssignal „TATAAA“ ist in blau dargestellt und befindet sich 69 bp upstream des poly(A)-Schwanzes. Die vorhandenen „UGUAN“-Motive sind in einer grünen Farbe eingezeichnet (Venkataraman *et al.*, 2005).

### 3.3.3 Die Mth<sub>550</sub>-cDNA besitzt ein Signalpeptid

Auch die Mth<sub>550</sub>-cDNA besitzt zu Beginn des ORF ein für Mollusken-Hämocyanine typisches Signalpeptid (Lieb *et al.*, 2001), welches hier eine Länge von 20 Aminosäuren (60 bp) aufweist. Zur Identifikation der Schnittstelle zwischen dem Signalpeptid und der 550 kDa Untereinheit wurde wiederum der „SIGNALP 3.0 SERVER“ (Bendtsen *et al.*, 2004) verwendet. Die graphische Ausgabe des Programms ist in Abb. 28 dargestellt.



**Abb. 28:** Graphische Ausgabe des „SIGNALP 3.0 SERVERS“ für die Mth<sub>550</sub>-Untereinheit.

Die Abbildung zeigt im unteren Bereich die Aminosäuresequenz der Mth<sub>550</sub>-Untereinheit beginnend mit dem Start-Methionin. Die X-Achse zeigt die Positionen der Sequenz an, während die Y-Achse die Höhe der einzelnen Scores angibt. Die Schnittstelle für die Signalpeptidase findet sich an der Stelle, wo der S-Score steil abfällt und der C-Score („cleavage site score“) signifikant hohe Werte erreicht. Zudem ist der Y-Score an der Schnittstelle, welche hier zwischen den Aminosäurepositionen 20 und 21 (SAA-AL) liegt, deutlich erhöht.

Das Signalpeptid der Mth<sub>550</sub>-Untereinheit beginnt mit der Aminosäure Methionin und endet mit der Abfolge Serin-Alanin-Alanin. Die folgende FU-a beginnt demnach an Position 21 in Abb. 28 mit den Aminosäuren Alanin und Leucin. Die Schnittstelle liegt genau an der Position, an der in der Grafik der S-Score steil abfällt und der C-Score („cleavage-site-score“) eine signifikante Größe erreicht. Auch der Y-Score, der eine Kombination aus dem S-Score und einem Derivat des C-Scores darstellt, bildet an der Schnittstelle einen signifikant hohen Gipfel.

### 3.3.4 Die Mth<sub>550</sub>-cDNA besitzt eine „Translation initiation site“

Auf die „Translation initiation sites“ wurde bereits im Abschnitt 3.2.4 eingegangen.

| -6 | -5 | -4 | -3 | -2 | -1 | START |   |   | +4 |
|----|----|----|----|----|----|-------|---|---|----|
| G  | T  | C  | A  | A  | A  | A     | T | G | A  |

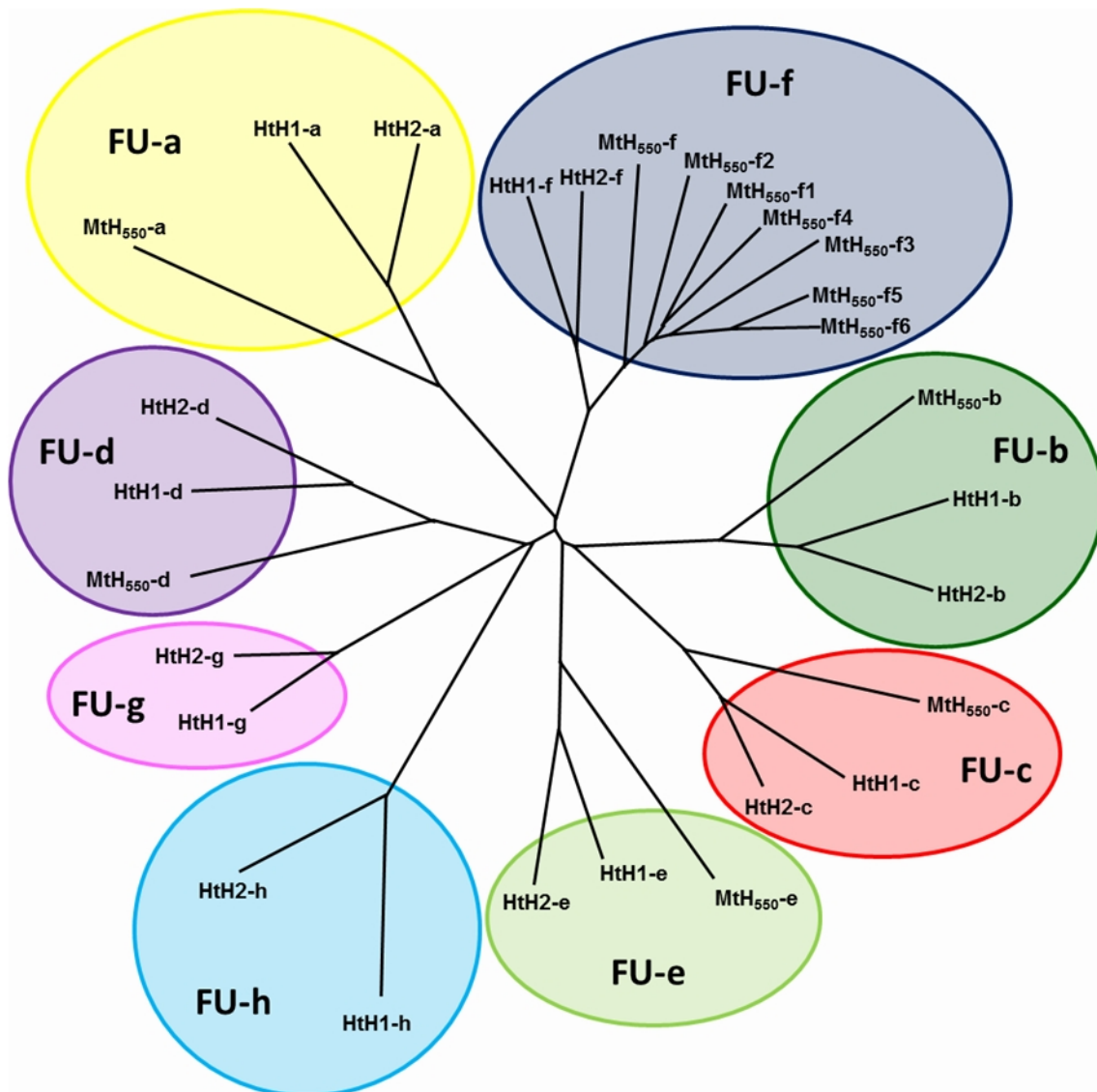
**Abb. 29:** TIS der Mth<sub>550</sub>-cDNA

Die Abbildung zeigt die TIS der für die Mth<sub>550</sub>-Untereinheit kodierenden cDNA von Position -6 bis +4. Das Startcodon „ATG“ ist rot hinterlegt. Die als wichtig für eine effektive Translation geltenden Purine an den Positionen -3 und +4 sind orange eingefärbt (Kozak, 1987).

Im gleichen Kapitel wird die TIS Consensus-Sequenz der Mollusken erläutert (Abb. 23). Die TIS der MtH<sub>550</sub>-cDNA ist in Abb. 29 dargestellt. Auch hier sind die für eine effektive Translation wichtigen Purine auf den Positionen -3 und +4 vorhanden (Kozak, 1987).

### 3.3.5 Die MtH<sub>550</sub>-Untereinheit besitzt sieben Varianten der FU-f

Die Untereinheit eines Gastropoden-Hämocyanins besteht typischerweise aus den acht paralogen funktionellen Einheiten FU-a bis FU-h (Lieb *et al.*, 1999; 2000).



**Abb. 30: Radialer Stammbaum der FUs des Hämocyanins von *Haliotis tuberculata* (HtH1 und HtH2) und der zwölf funktionellen Einheiten der MtH<sub>550</sub>-Untereinheit**

Aus den Aminosäuresequenzen der FUs wurde mit Hilfe des Programms „MUSCLE“ ein multiples Sequenzalignment erstellt. Darauf basierend wurde unter Verwendung des Neighbour-Joining Algorithmus ein phylogenetischer Stammbaum berechnet. Deutlich erkennbar ist die Gruppierung der orthologen FUs in jeweils einem Ast des Baumes (FU-a = gelb, FU-b = dunkelgrün, FU c = rot, FU d = violett, FU-e = hellgrün, FU-f = dunkelblau, FU-g = rosa, FU-h = hellblau).

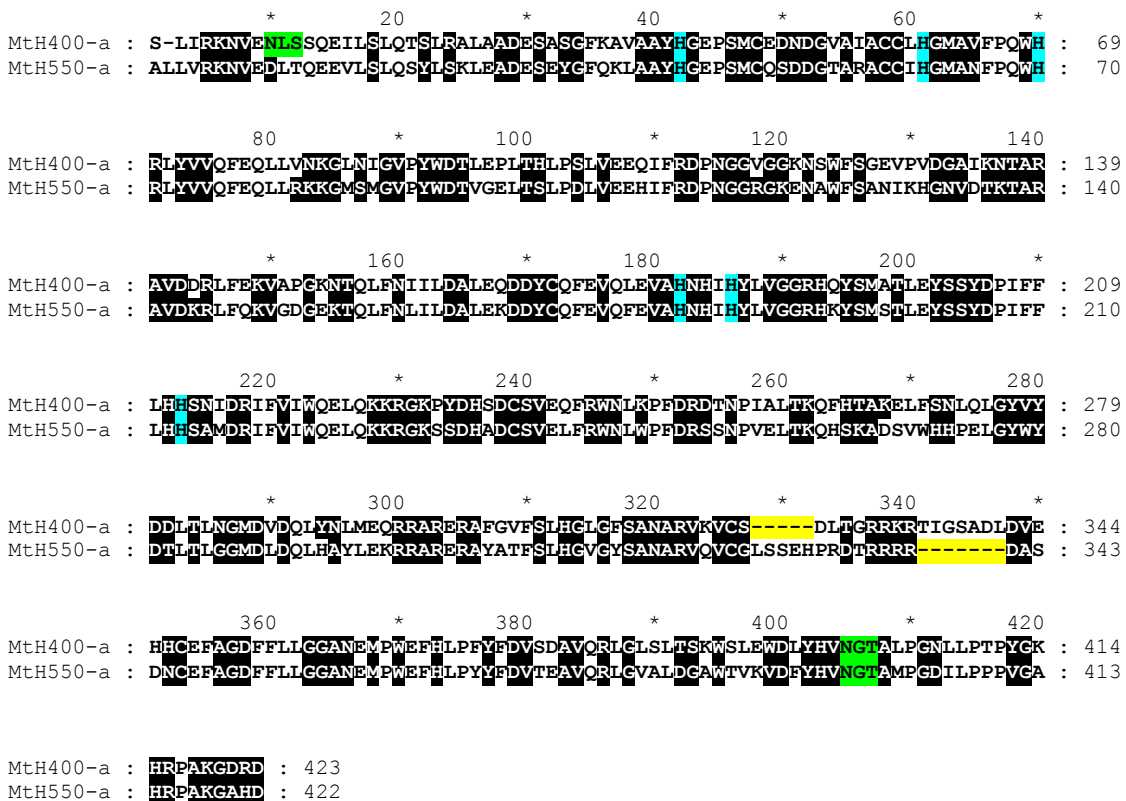
Die Sequenzdaten der über PCR-Reaktionen generierten DNA-Fragmente der MtH<sub>550</sub>-Untereinheit wurden unter Verwendung des Programms „MUSCLE“ mit den Daten des *Haliotis tuberculata* Hämocyanins aligniert. Aus dem Alignment wurde anschließend mit Hilfe des Programms „MEGA5“ ein radialer Stammbaum erstellt (siehe Abb. 30). Wie aus der Abb. 30 ersichtlich wird, sind die FU-a bis FU-f auch in der 550 kDa Untereinheit vollständig vorhanden. Im Gegensatz zur 400 kDa Untereinheit fehlen jedoch FU-g und FU-h, welche bei den typischen Gastropoden-Hämocyaninen für die Ausbildung des Kragenkomplexes verantwortlich sind (van Holde & Miller, 1995; Lieb *et al.*, 2000; Gatsogiannis & Markl, 2009). Stattdessen verfügt die MtH<sub>550</sub>-Untereinheit über weitere sechs Varianten der FU-f. Bei Betrachtung des Stammbaumes in Abb. 30 wird ersichtlich, dass alle sieben Varianten der FU-f der 550 kDa Untereinheit zusammen mit den beiden FU-f des HtH1 und HtH2 in einem Ast des Baumes gruppieren.

## 3.4 Vergleich der FUs von Mth<sub>400</sub> und Mth<sub>550</sub>

Die orthologen FUs der neu sequenzierten Untereinheiten des Mth wurden jeweils aligniert und analysiert.

### 3.4.1 FU-a

Die FU-a beider Untereinheiten (Mth<sub>400</sub>-a und Mth<sub>550</sub>-a) wurden aligniert und analysiert. Das resultierende Sequenzalignment zeigt Abb. 31. Beide FU-a besitzen die typischen hochkonservierten Histidine an den beiden Kupferbindungsstellen Cu-A und Cu-B. Die FU-a der 400 kDa Untereinheit verfügt über eine potentielle N-Glykosylierungsstelle („NLS“) an Position 10 des Alignments. Des Weiteren befindet sich bei beiden Mth-Untereinheiten eine N-Glykosylierungsstelle („NGT“) am C-Terminus der Aminosäuresequenz. Die beiden FU-a des Mth weisen eine Identität von 67 % zueinander auf.

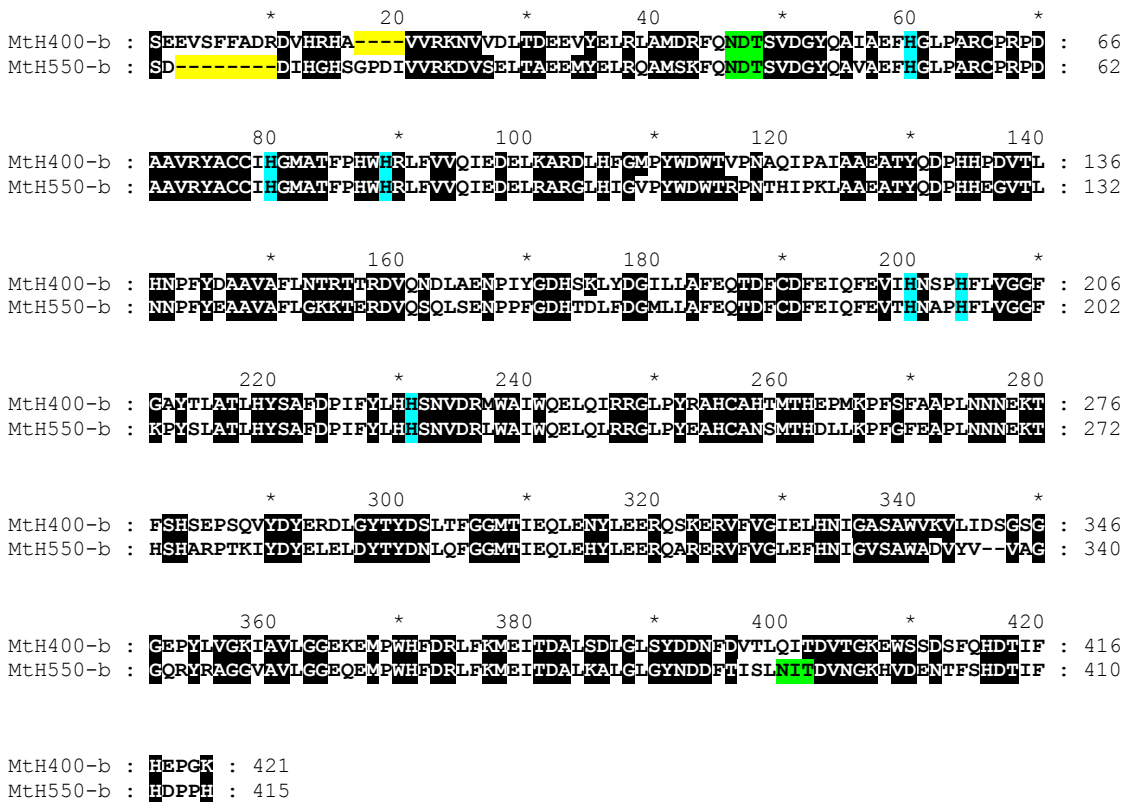


**Abb. 31:** Sequenzalignment der FU-a des Mth

Die Aminosäuresequenzen der FU-a des *Melanoides tuberculata* Hämocyanins (Mth<sub>400</sub>, Mth<sub>550</sub>) wurden mit Hilfe des Programmes „MAFFT“ und unter Verwendung der BLOSUM62 Matrix aligniert und das Alignment in „GENEDOC“ illustriert. Hochkonservierte Bereiche (100 %) sind schwarz hinterlegt. Die Histidine der beiden Kupferbindungsstellen (Cu-A und Cu-B) sind blau gefärbt. Die potentiellen N-Glykosylierungsstellen sind grün dargestellt. Die markanten Indels (gelb) werden unter 4.2 diskutiert.

### 3.4.2 FU-b

Das Sequenzalignment der beiden FU-b zeigt Abb. 32. Wie in der Abbildung zu erkennen ist, verfügen beide FU-b des MtH über die hochkonservierten Histidine der Kupferbindungsstellen. Potentielle N-Glykosylierungsstellen („NDT“) finden sich bei beiden MtH-Untereinheiten an Position 46 des Alignments. Eine weitere N-Glykosylierungsstelle („NIT“) besitzt die FU-b der MtH<sub>550</sub>-Untereinheit an Position 400. Die Identität der FUs-b der beiden MtH Untereinheiten liegt bei 72 %.



**Abb. 32:** Sequenzalignment der FU-b des MtH

Hochkonservierte Bereiche (100 %) sind schwarz hinterlegt. Die Histidine der beiden Kupferbindungsstellen (Cu-A und Cu-B) sind blau gefärbt. Die potentiellen N-Glykosylierungsstellen sind grün dargestellt. Die markanten Indels (gelb) im Linkerbereich werden unter 4.2 diskutiert.

### 3.4.3 FU-c

Die Sequenzen der FU-c des Mth zeigen deutlich, dass in beiden Untereinheiten N-Glykosylierungsstellen fehlen. Die sechs Histidine der beiden Kupferbindungsstellen sind bei beiden FU-c vorhanden. Die beiden FUs weisen eine Sequenzidentität von 70 % auf. Das entsprechende Alignment ist in Abb. 33 dargestellt.

```

Mth400-c : TFEVRC--VSDKRRRLDVDHLNSEQIQNLRDAFKALEDDTSINGENQIAAFHFGQPSWCPSFDAEIKKYACC : 68
Mth550-c : VHHADAGDAPSPDKVRSNIESLRTREQVQNLRDALQALEDDSSNAGFAQIAAFHFGEPWCPSFDAEIKKYACC : 70

Mth400-c : LHGMPTFFPWHRLITVQLENALRDKGFESGLPYWDWTLPMTELPESTFAETDNPVSCANBANPLHHAIV : 138
Mth550-c : LHGMVFFPWHRLITVQAEALRAHGFEGLPYWDWTRPMKALPQVVESAOVNVQQTGQTMNDNPFHHGIA : 140

Mth400-c : EGHKTRSVRGELFQOPAFGKFTIRIGEQVMLAFEQTDYCSFEVQFEVAFNHYIHALVGGNELYSMASLRYT : 208
Mth550-c : NGHATERDVSPKLFQOPEFGKFTDIAELVMFAFEQTDFCDFEVQMEIAHNYIHALVGGSKYYSMSLRYT : 210

Mth400-c : AYDPLFALHHSNTDRIWALWQALOKYRGLPYNTANCAITQLRKPLQPFQTSVTVNPDVSTRDHSVPFDVF : 278
Mth550-c : TYDPIFFLHHSNTDRIWALWQALOKYRGLPYNTANCALASMRKPLQPFQTSVTVNPDVSTRDHSVPFEIF : 280

Mth400-c : DYSSEFHYDYDNLEENGLSIAQLOHEVVVRKAHERVFAFEMLHGTKQSALVRFETCSSGCACCOAGEFYI : 348
Mth550-c : DYOSSFHYTYDDLEENGLSIPQLOHEVVVRQAHERVFAFEMLKGIGQSALVTFETCOGSGCCSAAGEFYM : 350

Mth400-c : LGDDMEPWFVYDRLEKYEITDQLKEQGLEPLDRYDIKYTVFALDGSVVGTDTEGQVTVVHTYGA : 412
Mth550-c : LGDDMEPWFVYDRLEKYEITDQLADKGLKPLDSYVVOYQVKDLONNVICTDNFGKVTIVHTYGA : 414

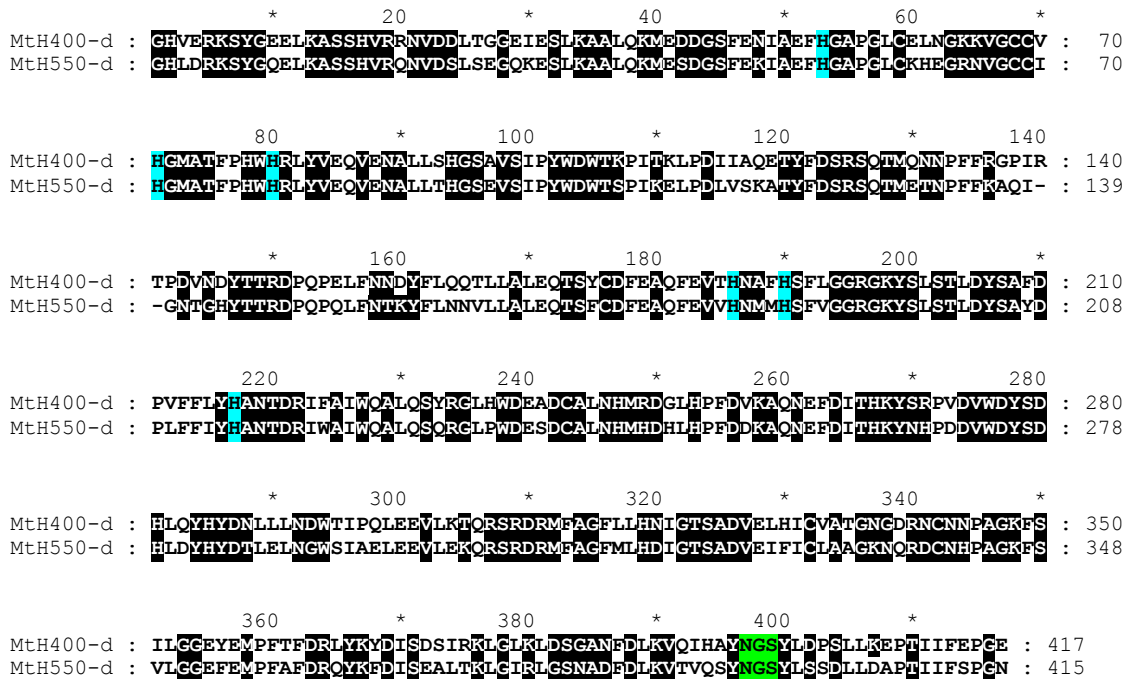
```

**Abb. 33:** Sequenzalignment der FU-c des Mth

Hochkonservierte Bereiche (100 %) sind schwarz hinterlegt. Die Histidine der beiden Kupferbindungsstellen (Cu-A und Cu-B) sind blau gefärbt.

### 3.4.4 FU-d

Auch das Alignment der beiden FU-d des Mth zeigt die hohe Konservierung der sechs kupferbindenden Histidine (siehe Abb. 34). Beide FU-d des Mth verfügen zudem über jeweils eine potentielle N-Glykosylierungsstelle an Position 397 des Alignments. Der Vergleich der beiden FU-d des *Melanoides tuberculata* Hämocyanins ergibt eine Sequenzidentität von 74 %.

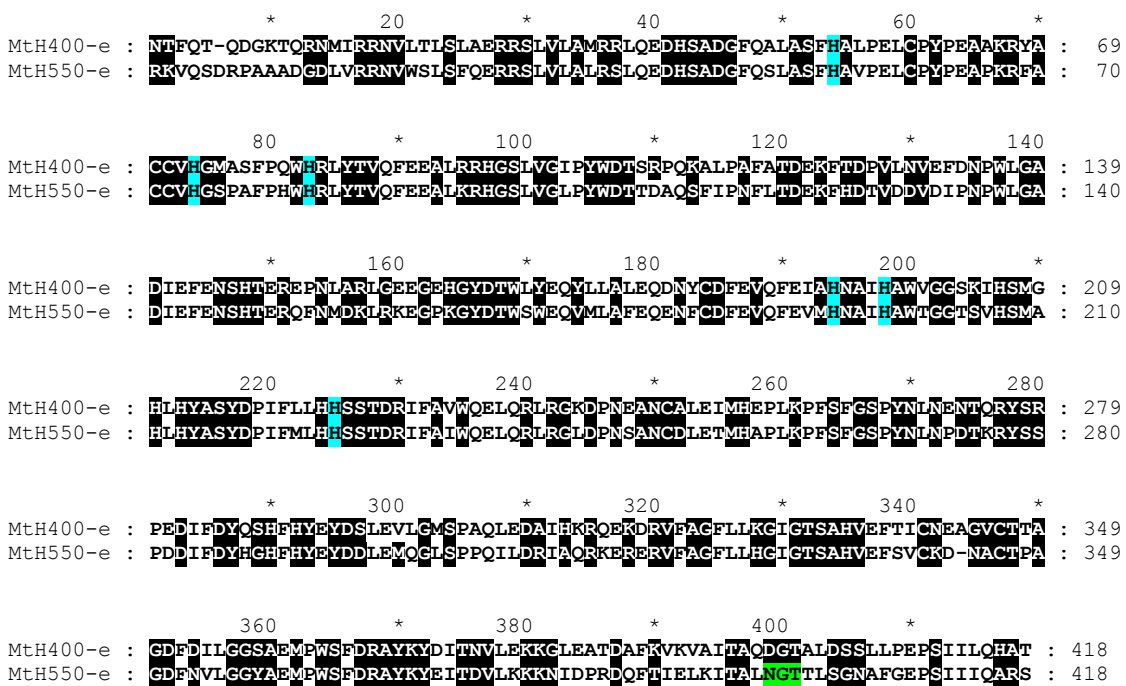


**Abb. 34:** Sequenzalignment der FU-d des Mth

Hochkonservierte Bereiche (100 %) sind schwarz hinterlegt. Die Histidine der beiden Kupferbindungsstellen (Cu-A und Cu-B) sind blau gefärbt. Die potentiellen N-Glykosylierungsstellen sind grün dargestellt.

### 3.4.5 FU-e

Das Sequenzalignment der beiden FU-e der 400 kDa sowie der 550 kDa Untereinheit des Mth ist in Abb. 35 dargestellt. Die sechs Histidine der Kupferbindungsstellen sind bei beiden in das Alignment einbezogenen FUs wie erwartet vorhanden. Die FU-e der 400 kDa Untereinheit des Mth besitzt keine potentielle N-Glykosylierungsstelle. Bei der FU-e der Mth<sub>550</sub>-Untereinheit hingegen befindet sich eine N-Glykosylierungsstelle („NGT“) an Position 399 des Alignments. Der direkte Sequenzvergleich der beiden FUs der Mth Untereinheiten ergibt eine Sequenzidentität von 68 %.

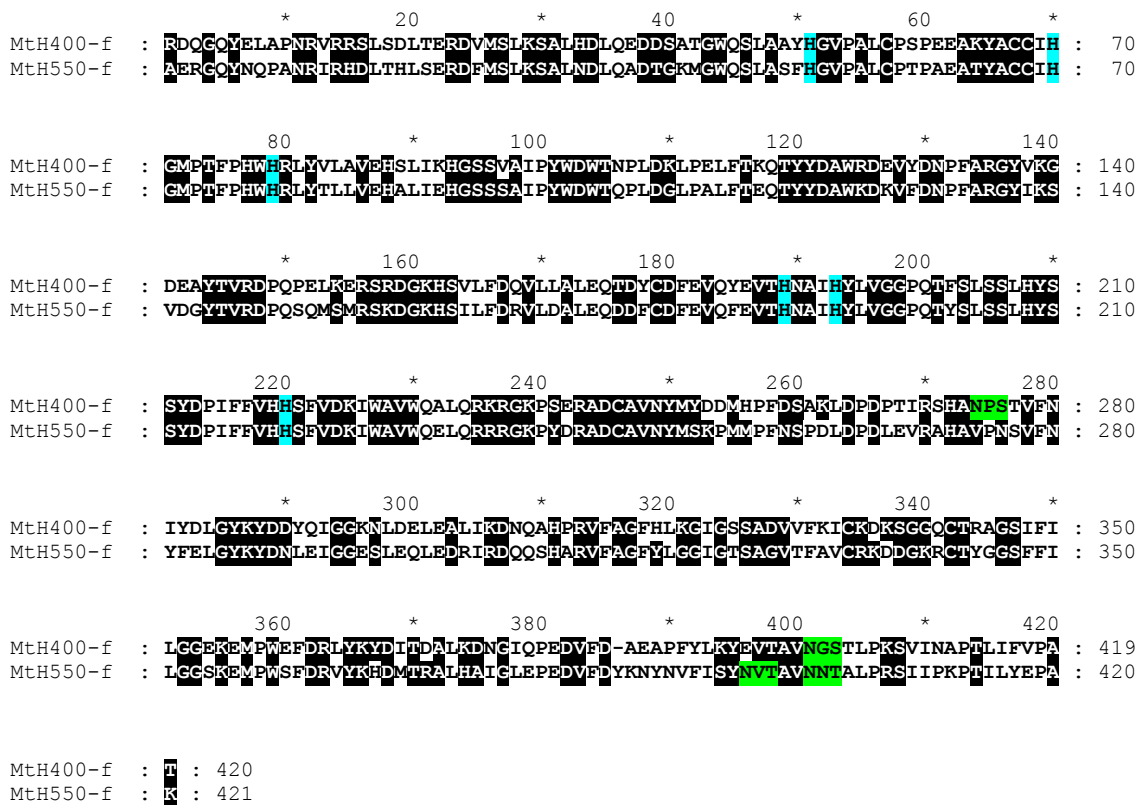


**Abb. 35:** Sequenzalignment der FU-e des Mth

Hochkonservierte Bereiche (100 %) sind schwarz hinterlegt. Die Histidine der beiden Kupferbindungsstellen (Cu-A und Cu-B) sind blau gefärbt. Die potentiellen N-Glykosylierungsstellen sind grün dargestellt.

### 3.4.6 FU-f

Abb. 36 zeigt das Alignment der in der Zylinderwand gelegenen FU-f beider Untereinheiten des MtH. Die FU-f der 400 kDa Untereinheit des MtH verfügt über eine potentielle N-Glykosylierungsstelle („NGS“) an Position 401 des Alignments. Ein weiteres NXS-Sequenzmotiv („NPS“) befindet sich an Position 274. Motive dieser Art, die ein Prolin in der „X“-Position aufweisen, scheinen die Glykosylierung an der entsprechenden Stelle vollständig zu blockieren (Bause, 1983; Shakin-Eshleman *et al.*, 1996). Die FU-f der 550 kDa Untereinheit besitzt zwei NXT-Sequenzmotive („NVT“ sowie „NNT“) die an den Alignmentpositionen 396 und 401 liegen. Die sechs kupferbindenden Histidine sind in beiden alignierten FU-f vollständig enthalten. Die Identität der beiden FU-f des *Melanoides tuberculata* Hämocyanins beträgt 67 %.



**Abb. 36:** Sequenzalignment der in der Zylinderwand gelegenen FU-f der beiden MtH-Untereinheiten. Hochkonservierte Bereiche (100 %) sind schwarz hinterlegt. Die Histidine der beiden Kupferbindungsstellen (Cu-A und Cu-B) sind blau gefärbt. Die potentiellen N-Glykosylierungsstellen sind grün dargestellt.

### 3.4.7 FU-g

Die Sequenz der FU-g der Mth<sub>400</sub>-Untereinheit des *Melanoides tuberculata* Hämocyanins ist in Abb. 37 dargestellt. Deutlich erkennbar verfügt auch die FU-g des Mth über die sechs hochkonservierten kupferbindenden Histidine sowie über eine potentielle N-Glykosylierungsstelle („NGT“) an Position 383 der Sequenz. Die für Mollusken-Hämocyanine typische Verkürzung der FU-g (Lieb & Todt, 2008) im Vergleich zu den anderen paralogen FUs ist auch beim Mth zu beobachten. Mit einer Länge von 403 bp ist die FU-g bis zu 25 Aminosäuren kürzer als die paralogen FU-a bis FU-f.

```

Mth400-g :   *           20           *           40           *           60           *
           : GASVDAKSYAVAGMSVRKDINTLSSAETQSLKNSLQKVM DGEGLPSYQKIAEMHGWPGDCEQNGKVVACC : 70

Mth400-g :           80           *           100           *           120           *           140
           : HGMASFPQWHRVYTRLLEMAMTWQGANVGIPYWDWTEAFTALPSLVTDDGDNPFNHGHIEALGKVSRE : 140

Mth400-g :           *           160           *           180           *           200           *
           : PRPQLFKDPEHGEESFFYRQILLAFEQRDYCDFEVQFEVTHNAPHHSWIGGTSPYGMSTLEFSSYDPVFFL : 210

Mth400-g :           220           *           240           *           260           *           280
           : HHSNVDRQFAIWQALQKYRGLDYNSANCDIQELRMDLEPFNRDYIPIVIIRNNARAIDAFNYDQYSYQYD : 280

Mth400-g :           *           300           *           320           *           340           *
           : NLNFHGMSIPELEALLEKRRESDHVFLNFMLLGGASADVTFDIDCDSEGHCFAGTFFAVLGGPLEMPWSF : 350

Mth400-g :           360           *           380           *           400
           : DRLFKYDVTSVFRQMHLRPDSDYSVKVHIVSVNGTELDSDLLEAPSVS FVPGK : 403

```

**Abb. 37: Sequenz der FU-g der Mth<sub>400</sub>-Untereinheit**

Die FU-g ist üblicherweise am Kragenkomplex beteiligt (Gatsogiannis & Markl, 2009). Die Histidine der beiden Kupferbindungsstellen (Cu-A und Cu-B) sind blau gefärbt. Die potentielle N-Glykosylierungsstelle ist grün dargestellt.

### 3.4.8 FU-h

Die Sequenz der FU-h des Mth<sub>400</sub> ist in Abb. 38 dargestellt.

```

Mth400-h : RGRKAREDYRPTVSEDLIRKEVSSLSLEEVSNLKNALYKLNQNDHGPNGF*EAIASF*H*GDPGLCP*EGSGSKY : 70
                *           20           *           40           *           60           *
Mth400-h : ACCQ*H*GMPV*FPHW*H*RLLTVQ*FERAL*KAKGSV*VGVPYWDW*TRP*SRGI*PQLF*SNTY*GNNP*FLTYEI*QAASAF : 140
                *           80           *           100          *           120          *           140
Mth400-h : VERDVQ*VDAL*EHTT*GHEGF*EDSL*FHQALE*TLEETS*YCEFEV*QFEM*LHNAV*H*ALVGG*TKTH*SMATLEWSAF : 210
                *           160          *           180          *           200          *
Mth400-h : DPFFM*VH*HS*SIDRI*WRIW*QELQ*RLR*KKPF*NVARC*AMRYL*RKPLE*PFSYAS*VNTDE*VARTNS*RPID*IFDTA : 280
                *           220          *           240          *           260          *           280
Mth400-h : KFH*YDFD*NLDL*GGHS*IGEV*NKMIN*DMRS*KTRL*FVGF*VLSG*IETS*SATV*KIEM*DDSQ*GNSHE*VGTFFV*LVGGH : 350
                *           300          *           320          *           340          *
Mth400-h : NEMP*WAYER*VYKYE*ITEAA*KYGL*DHDS*VNF*FHHK*VVKYD*GQEL*SAQ*FKAP*ILVER*PAG*VDYD*VAI*FQ*LK : 420
                *           360          *           380          *           400          *           420
Mth400-h : KDND*LAPK*VVL*PRG*TRIR*FHSA*EDGV*SSV*LREMG*SYTN*AVYCS*I*PPGD*ANTY*KLEVE*Y*TLEP*GDY*FFVSN : 490
                *           440          *           460          *           480          *
Mth400-h : DEEK*CKQ*GDRI*QIS*IDEE : 508
                *           500

```

#### Abb. 38: Sequenz der FU-h der Mth<sub>400</sub>-Untereinheit

Die FU-h ist üblicherweise am Kragenkomplex beteiligt (Gatsogiannis & Markl, 2009). Die Histidine der beiden Kupferbindungsstellen (Cu-A und Cu-B) sind blau gefärbt. Der 98 Aminosäuren lange „Tail“ der FU-h ist rot hinterlegt.

Deutlich erkennbar ist, dass die FU-h über keine potentiellen N-Glykosylierungsstellen verfügt. Die sechs kupferbindenden Histidine sind vollständig vorhanden. Der für die FU-h von Mollusken-Hämocyaninen typische C-terminale „Schwanz“ von ~100 Aminosäuren Länge (Keller *et al.*, 1999; Lieb *et al.*, 2000) existiert auch bei der FU-h der Mth<sub>400</sub>-Untereinheit.

### 3.4.9 FU-f<sub>1</sub> bis FU-f<sub>6</sub>

Die 550 kDa Untereinheit des Mth verfügt über sechs zusätzliche C-terminale Versionen der FU-f (siehe 3.3.5). Diese wurden zusammen mit den beiden an der Bildung der Zylinderwand beteiligten („internen“) FU-f der 400 kDa und 550 kDa Untereinheit aligniert. Das resultierende multiple Sequenzalignment zeigt Abb. 39. Alle FU-f des Mth verfügen über die sechs hochkonservierten Histidine, welche für die Bindung des Kupfers im aktiven Zentrum erforderlich sind. FU-f<sub>1</sub>, FU-f<sub>2</sub> und FU-f<sub>4</sub> besitzen jeweils eine potentielle N-Glykosylierungsstelle („NGS“) an Alignmentposition 409. An der entsprechenden Position liegen auch N-Glykosylierungsstellen der „eigentlichen“ FU-f des Mth.

**Tab.1: Identitätsmatrix der FU-f des *Melanoides tuberculata* Hämocyanins**

Alle FU-f Varianten wurden miteinander verglichen und eine Identitätsmatrix aufgestellt. Die Sequenzpaare mit den größten Übereinstimmungen sind rot hinterlegt.

| Identität [%]                      | MtH <sub>400</sub> -f | MtH <sub>550</sub> -f | MtH <sub>550</sub> -f <sub>1</sub> | MtH <sub>550</sub> -f <sub>2</sub> | MtH <sub>550</sub> -f <sub>3</sub> | MtH <sub>550</sub> -f <sub>4</sub> | MtH <sub>550</sub> -f <sub>5</sub> | MtH <sub>550</sub> -f <sub>6</sub> |
|------------------------------------|-----------------------|-----------------------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|
| MtH <sub>550</sub> -f              | 67 %                  |                       |                                    |                                    |                                    |                                    |                                    |                                    |
| MtH <sub>550</sub> -f <sub>1</sub> | 59 %                  | 60 %                  |                                    |                                    |                                    |                                    |                                    |                                    |
| MtH <sub>550</sub> -f <sub>2</sub> | 57 %                  | 56 %                  | 62 %                               |                                    |                                    |                                    |                                    |                                    |
| MtH <sub>550</sub> -f <sub>3</sub> | 54 %                  | 55 %                  | 62 %                               | 60 %                               |                                    |                                    |                                    |                                    |
| MtH <sub>550</sub> -f <sub>4</sub> | 57 %                  | 57 %                  | 66 %                               | 63 %                               | 59 %                               |                                    |                                    |                                    |
| MtH <sub>550</sub> -f <sub>5</sub> | 56 %                  | 58 %                  | 62 %                               | 60 %                               | 63 %                               | 66 %                               |                                    |                                    |
| MtH <sub>550</sub> -f <sub>6</sub> | 55 %                  | 58 %                  | 63 %                               | 59 %                               | 62 %                               | 66 %                               | 72 %                               |                                    |

Bei den FU-f<sub>5</sub> und FU-f<sub>6</sub> finden sich die Sequenzmotive für die N-Glykosylierung an Alignmentposition 139 („NES“). Die FU-f<sub>5</sub> verfügt zusätzlich über ein prolinhaltiges und damit die N-Glykosylierung hemmendes (Bause, 1983; Shakin-Eshleman *et al.*, 1996) „NXS“-Motiv an Position 387 („NPS“). Die FU-f<sub>3</sub> verfügt über kein potentielles Sequenzmotiv für die N-Glykosylierung. Insgesamt sind somit fünf N-Glykosylierungsstellen in den sechs C-terminalen FU-f der MtH<sub>550</sub>-Untereinheit vorhanden.

Auffällig ist außerdem die Verkürzung der Peptidlinker am N-Terminus der FU-f<sub>1</sub> bis FU-f<sub>6</sub>, wobei im Vergleich zu den „eigentlichen“ FU-f jeweils zwischen vier und zehn Aminosäuren deletiert sind.

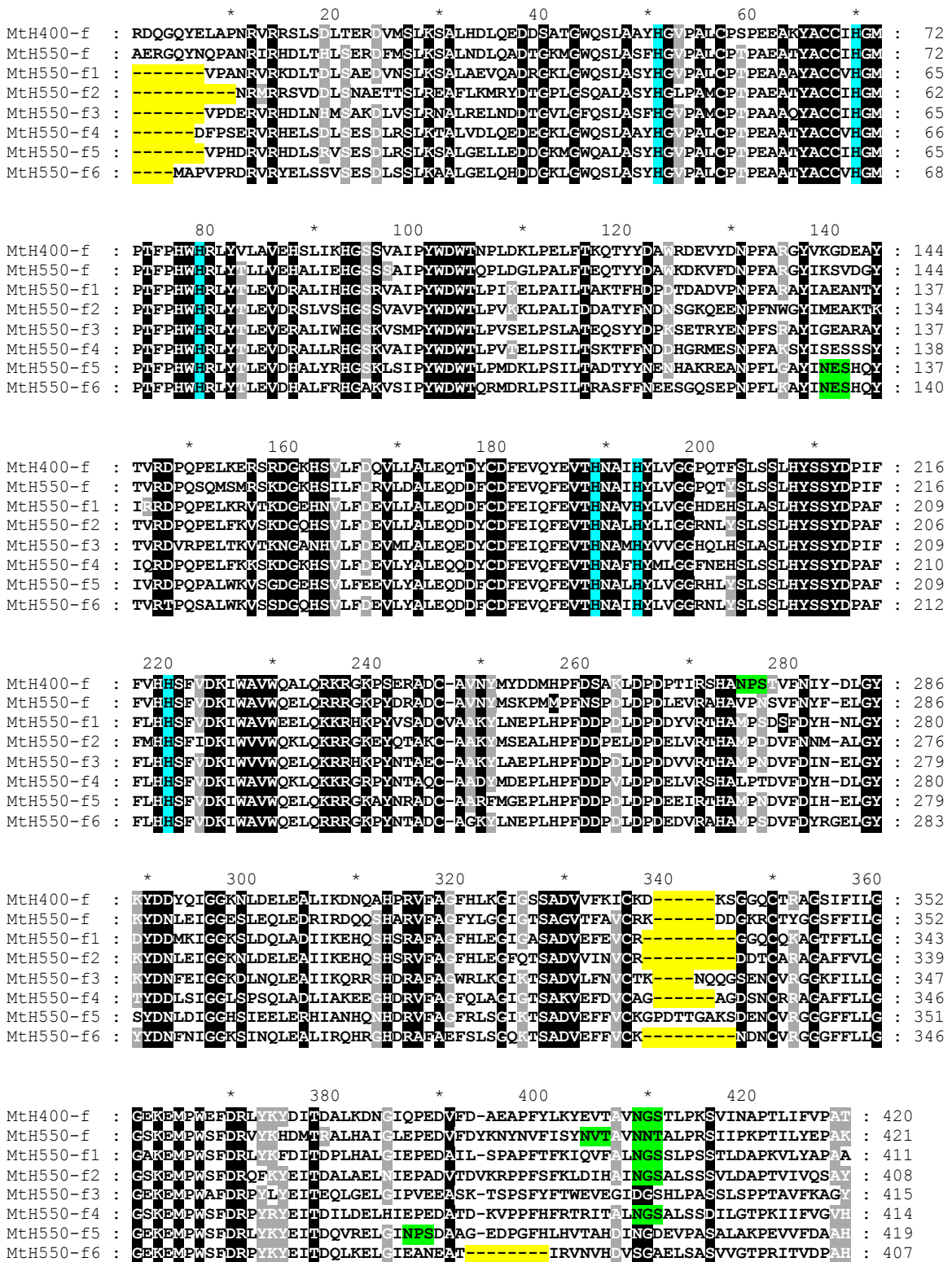


Abb. 39: Multiples Sequenzalignment der FU-f des Mth

Die Aminosäuresequenzen der FU-f Varianten des *Melanoides tuberculata* Hämocyanins (Mth<sub>400</sub>, Mth<sub>550</sub>) wurden mit Hilfe des Programmes „MAFFT“ und unter Verwendung der BLOSUM62 Matrix aligniert und das Alignment in „GENEDOC“ illustriert. Hochkonservierte Bereiche (100 %) sind schwarz, konservierte Bereiche (>90 %) grau hinterlegt. Die Histidine der beiden Kupferbindungsstellen sind blau gefärbt. Die potentiellen N-Glykosylierungsstellen sind grün dargestellt. Die markanten Indels (gelb) werden unter 4.2 diskutiert.

Die acht FU-f Varianten des MtH wurden paarweise miteinander verglichen und so die Identität zwischen den Sequenzpaaren bestimmt. In Tab.1 ist die aus den Vergleichen resultierende Identitätsmatrix dargestellt. Aus der Tabelle wird ersichtlich, dass die „eigentlichen“ FU-f der beiden MtH Untereinheiten die zweithöchste Identität aufweisen (67 %). Noch ähnlicher sind sich nur die beiden letzten FUs der MtH<sub>550</sub>-Untereinheit - nämlich FU-f<sub>5</sub> und FU-f<sub>6</sub> (72 %). Darüber hinaus liegen die Identitäten zwischen den FU-f des MtH zwischen 54 % und 66 %.

### 3.4.10 Statistische Daten der Untereinheiten des MtH

Die ermittelten cDNA-Sequenzen des MtH wurden in die abgeleitete Aminosäuresequenz übersetzt und analysiert. Dabei wurde sowohl der theoretische isoelektrische Punkt als auch die theoretische Molekularmasse in kDa bestimmt. Tabelle 2 fasst diese Daten sowie die Anzahl der potentiellen N-Glykosylierungsstellen und die Länge der einzelnen funktionellen Einheiten in Aminosäuren zusammen. Die durchschnittliche theoretische Molekularmasse der FUs des MtH beträgt ca. 48 kDa. Da je nach Glykosylierungsstatus der FUs zusätzlich 1000 bis 2000 Da für die Kohlenhydratseitenketten dazu addiert werden müssen (Miller *et al.*, 1998) entsprechen die funktionellen Einheiten des Hämocyanins von *Melanoides tuberculata* der für andere Gastropoden-Hämocyanine beschriebenen Molekularmasse von ca. 50 kDa/FU (van Holde & Miller, 1995; Gebauer *et al.*, 1994; Lieb *et al.*, 2000).

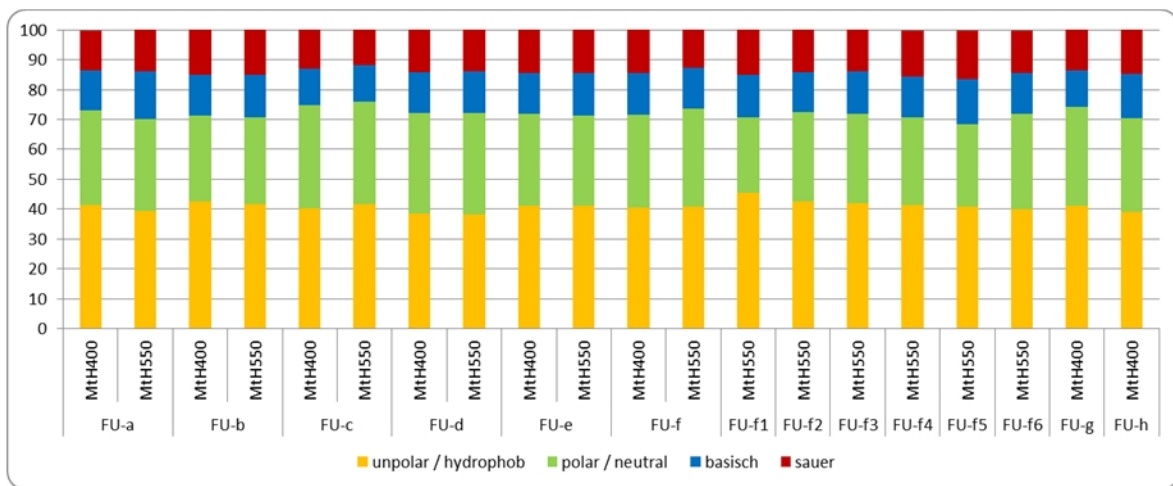
Der durchschnittliche isoelektrische Punkt der FUs des MtH liegt bei einem pH-Wert von 5,33. Mit Ausnahme der FU-c, FU-e (MtH<sub>400</sub>), FU-h sowie der FU-f<sub>3</sub> besitzen alle FUs eine bzw. zwei N-Glykosylierungsstellen. Insgesamt verfügt die 550 kDa Untereinheit über zwölf und die 400 kDa Untereinheit über sechs potentielle N-Glykosylierungsstellen.

Tab.2: Statistische Daten der Untereinheiten des Hämocyanins von *Melanoides tuberculata*

| Struktur                 | Aminosäuren | pI          | Molekularmasse<br>[kDa] | potentielle<br>N-Glykosylierungsstellen |
|--------------------------|-------------|-------------|-------------------------|---|
| <b>MtH<sub>400</sub></b> | <b>3422</b> | <b>5,16</b> | <b>390,1</b>            | <b>6</b>                                |
| Signalpeptid             | 23          |             |                         |   |
| FU-a                     | 423         | 5,91        | 47,98                   | 2                                       |
| FU-b                     | 421         | 4,77        | 48,35                   | 1                                       |
| FU-c                     | 412         | 5,04        | 46,89                   | 0                                       |
| FU-d                     | 417         | 5,07        | 47,79                   | 1                                       |
| FU-e                     | 418         | 5,1         | 47,67                   | 0                                       |
| FU-f                     | 420         | 5,28        | 47,84                   | 1                                       |
| FU-g                     | 403         | 4,82        | 45,81                   | 1                                       |
| FU-h                     | 508         | 5,77        | 57,89                   | 0                                       |
| <b>MtH<sub>550</sub></b> | <b>4979</b> | <b>5,38</b> | <b>566,78</b>           | <b>12</b>                               |
| Signalpeptid             | 20          |             |                         |   |
| FU-a                     | 422         | 6,36        | 48,28                   | 1                                       |
| FU-b                     | 415         | 5,15        | 47,55                   | 2                                       |
| FU-c                     | 414         | 5,47        | 47,16                   | 0                                       |
| FU-d                     | 415         | 5,25        | 47,41                   | 1                                       |
| FU-e                     | 418         | 5,49        | 47,93                   | 1                                       |
| FU-f                     | 421         | 6,13        | 48,13                   | 2                                       |
| FU-f <sub>1</sub>        | 411         | 5,06        | 46,1                    | 1                                       |
| FU-f <sub>2</sub>        | 408         | 5,16        | 46,32                   | 1                                       |
| FU-f <sub>3</sub>        | 415         | 5,63        | 47,4                    | 0                                       |
| FU-f <sub>4</sub>        | 414         | 4,93        | 46,86                   | 1                                       |
| FU-f <sub>5</sub>        | 419         | 5,02        | 47,56                   | 1                                       |
| FU-f <sub>6</sub>        | 407         | 5,26        | 46,28                   | 1                                       |

### 3.5 Aminosäurezusammensetzung des MtH

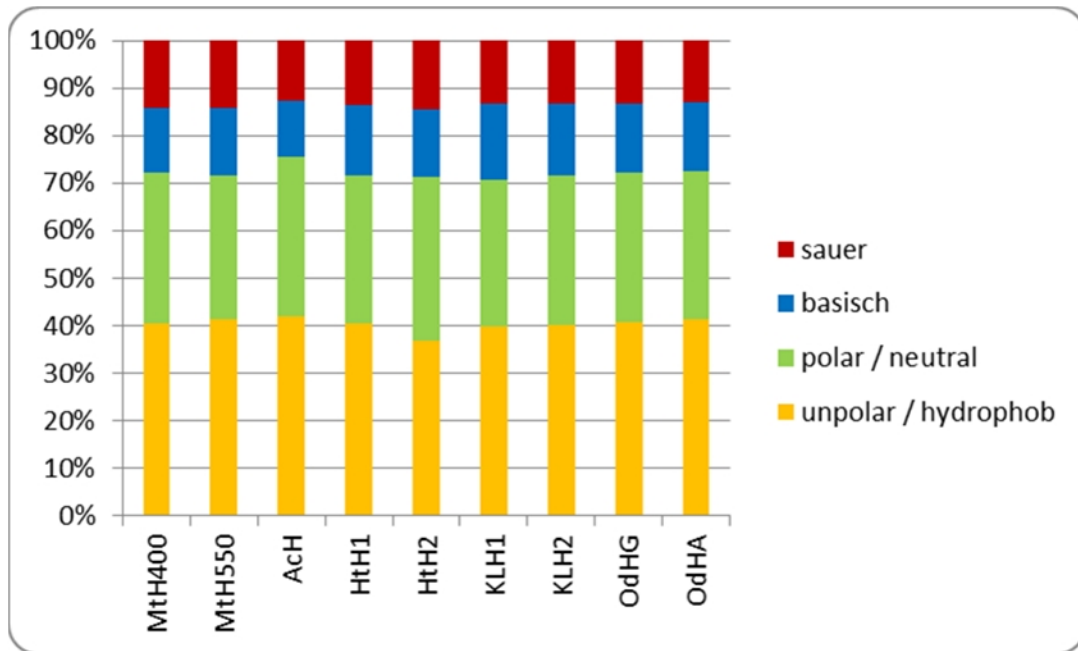
Die Sequenzdaten des MtH wurden in die abgeleitete Aminosäuresequenz übersetzt und mit Hilfe des Programms „GENEIOUS“ auf ihre Aminosäurezusammensetzung hin untersucht. In Abb. 40 ist die prozentuale Aminosäurezusammensetzung der funktionellen Einheiten des MtH grafisch dargestellt. Dabei wurden die 20 proteinogenen Aminosäuren anhand ihrer physikalisch-chemischen Eigenschaften in die vier Gruppen unpolar/hydrophob, polar/neutral, basisch sowie sauer zusammengefasst. Die Abbildung lässt deutlich erkennen, dass sich die einzelnen funktionellen Einheiten hinsichtlich des Verhältnisses der vier Aminosäuregruppen nicht merklich unterscheiden.



**Abb. 40: Prozentuale Aminosäurezusammensetzung der FUs des MtH**

Die 20 proteinogenen Aminosäuren wurden anhand ihrer physikalisch-chemischen Eigenschaften in vier Gruppen zusammengefasst: unpolar/hydrophob (Ala, Ile, Leu, Met, Phe, Pro, Trp, Val), polar/neutral (Asp, Cys, Gln, Gly, Ser, Thr, Tyr), basisch (Arg, His, Lys) und sauer (Asp, Gln). Das Diagramm zeigt den prozentualen Anteil der einzelnen Aminosäuregruppen in den verschiedenen funktionellen Einheiten des MtH.

Abb. 41 zeigt die Aminosäurezusammensetzung der beiden kompletten Untereinheiten des MtH. Auch auf Ebene der Untereinheiten lassen sich sowohl zwischen MtH<sub>400</sub> und MtH<sub>550</sub> als auch zwischen dem *Melanoides tuberculata* Hämocyanin und anderen Mollusken-Hämocyaninen keine signifikanten Unterschiede in der Aminosäurezusammensetzung feststellen. Die vollständige tabellarische Aminosäureverteilung in den Untereinheiten des MtH befindet sich im Anhang der vorliegenden Arbeit (I).



**Abb. 41: Prozentuale Aminosäurezusammensetzung verschiedener Mollusken-Hämocyanine**

Die 20 proteinogenen Aminosäuren wurden anhand ihrer physikalisch-chemischen Eigenschaften in vier Gruppen zusammengefasst: unpolar/hydrophob (Ala, Ile, Leu, Met, Phe, Pro, Trp, Val), polar/neutral (Asp, Cys, Gln, Gly, Ser, Thr, Tyr), basisch (Arg, His, Lys) und sauer (Asp, Gln). Das Diagramm zeigt den prozentualen Anteil der einzelnen Aminosäuregruppen in den Untereinheiten der Hämocyanine von *Melanoides tuberculata* (MtH<sub>400</sub>, MtH<sub>550</sub>), *Aplysia californica* (AcH), *Haliotis tuberculata* (HtH1, HtH2), *Megathura crenulata* (KLH1, KLH2) sowie *Octopus dofleini* (OdH<sub>G</sub>, OdH<sub>A</sub>).

## 4 Diskussion

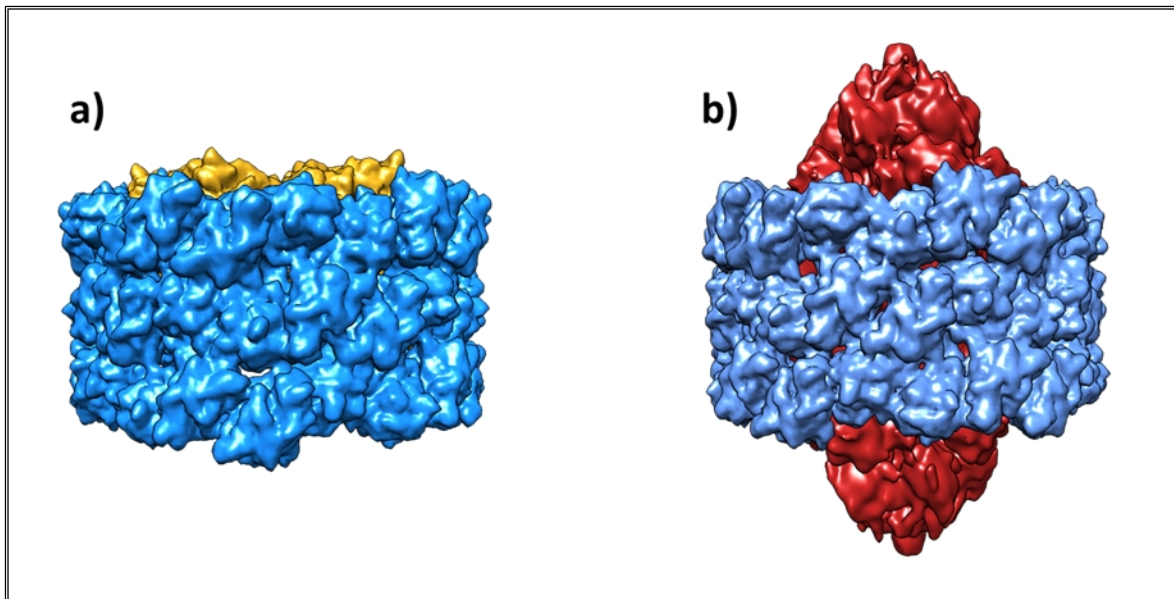
Durch die proteinbiochemischen und elektronenmikroskopischen Untersuchungen von Lieb *et al.* (2010) war bekannt, dass sich das Mega-Hämocyanin der Schnecke *Melanoides tuberculata* aus einer 400 kDa sowie einer 550 kDa großen Untereinheit zusammensetzt. Weitergehende Versuche konnten zudem zeigen, dass sich beide Untereinheiten immunologisch deutlich voneinander unterscheiden (Lieb *et al.*, 2010).

Das vorrangige Ziel dieser Arbeit war es, die cDNA-Sequenz und somit auch die Primärstruktur des Mega-Hämocyanins von *Melanoides tuberculata* zu ermitteln. Es gelang zwei deutlich verschiedene cDNAs vollständig zu sequenzieren. Dieses Ergebnis deckt sich mit den beiden von Lieb *et al.* (2010) proteinbiochemisch charakterisierten Polypeptiden, bei welchen es sich um die beiden Untereinheiten des Mega-Hämocyanins handelt.

Die 10915 bp umfassende cDNA der MtH<sub>400</sub>-Untereinheit entspricht strukturell einem typischen Gastropoden-Hämocyanin. Der kodierende Bereich umfasst ein Signalpeptid sowie die für Gastropoden-Hämocyanine typischen FU-a bis FU-h (Lieb *et al.*, 1999). Anhand von Reassoziationsversuchen konnten Lieb *et al.* (2010) zeigen, dass die 400 kDa Untereinheit zu typischen Dekameren mit einseitigem Kragenkomplex assembliert (siehe Abb. 42a). Diese Dekamere bildeten auch höhere Organisationsformen wie Di- und Multidekamere aus. Zusammen mit den neuen Primärstrukturdaten zeigen diese Beobachtungen deutlich, dass es sich bei der MtH<sub>400</sub>-Untereinheit des *Melanoides tuberculata* Hämocyanins um die Untereinheit eines typischen Gastropoden-Hämocyanins handelt.

Die mit 15346 bp deutlich längere cDNA der 550 kDa Untereinheit weicht in der Primärstruktur des kodierenden Bereichs auffallend von einem typischen Gastropoden-Hämocyanin ab. Die an der Bildung der Zylinderwand beteiligten FU-a bis FU-f sind in der MtH<sub>550</sub>-Untereinheit vollständig vorhanden. Diese Daten decken sich mit den elektronenmikroskopischen Befunden von Lieb *et al.* (2010), die in der Wandstruktur des Mega-Dekamers eine typische, aus den FU-a bis FU-f bestehende Wandstruktur eines Gastropoden-Hämocyanins sahen (Abb. 42b). Auf die FU-f folgen jedoch nicht die typischerweise vorhandenen FU-g und FU-h sondern sechs weitere Varianten der FU-f (FU-f<sub>1</sub> bis FU-f<sub>6</sub>). Diese höchst ungewöhnliche Struktur konnte bisher bei keinem anderen untersuchten Mollusken-Hämocyanin beobachtet werden. Abbildung 42b zeigt die

Seitenansicht einer von Lieb *et al.* (2010) veröffentlichten 3D-Rekonstruktion eines Mega-Dekamers mit einer Auflösung von 13 Å. Erkennbar sind die Zylinderwand (blau) sowie der im Vergleich zu typischen Hämocyanin-Dekameren deutlich abgewandelte Kragenkomplex (rot). Dieser Kragenkomplex füllt den Zylinder fast vollständig aus und bildet zudem auf beiden Seiten des Mega-Dekamers annähernd kegelförmige Erweiterungen (Abb. 42b).



**Abb. 42:** 3D-Rekonstruktion des Mega-Hämocyanins von *Leptoxis carinata*

3D-Rekonstruktionen des Dekamers (a) sowie des Mega-Dekamers (b) des Hämocyanins von *Leptoxis carinata* (Cerithioidea) mit einer Auflösung von 13 Å. Deutlich erkennbar sind die beiden ähnlichen Zylinderwände in blauer Farbe sowie die unterschiedlichen Kragenkomplexe, die in einem gelben bzw. roten Farbton dargestellt sind (nach Lieb *et al.*, 2010).

In den Reassoziationsversuchen von Lieb *et al.* (2010) wurde festgestellt, dass die isolierte 550 kDa Untereinheit des *Melanoides tuberculata* Hämocyanins ausschließlich Mega-Dekamere bildete. Da die Zylinderwand augenscheinlich der eines typischen Gastropoden-Hämocyanins entspricht, sollten an ihrer Bildung die FU-a bis FU-f beteiligt sein. Die in Abbildung 42b rotgefärbte Proteinmasse muss demnach durch die FU-f<sub>1</sub> bis FU-f<sub>6</sub> gebildet werden. Lieb *et al.* (2010) gehen davon aus, dass der Kragenkomplex des Mega-Dekamers durch zehn Kopien eines etwa 250 kDa großen bzw. ca. 2000 Aminosäuren umfassenden Anteils der 550 kDa Untereinheit gebildet wird. Dies entspräche laut den Schätzungen von Lieb *et al.* (2010) etwa vier bis fünf zusätzlichen funktionellen Einheiten. Durch die im Rahmen dieser Arbeit ermittelten Primärstrukturdaten der 550 kDa Untereinheit wird klar, dass der Kragenkomplex des Mega-Dekamers durch insgesamt sechs zusätzliche funktionelle Einheiten gebildet wird. Sie umfassen insgesamt 2474 Aminosäuren und

haben eine theoretische Molekularmasse von 280 kDa. Die Genauigkeit der Extrapolationsmethode von Lieb *et al.* (2010) ist erstaunlich. Die geschätzten 550 kDa stimmen auffallend gut mit den anhand der Primärstruktur ermittelten 567 kDa (ohne Kohlenhydratseitenketten) überein.

Alle sechs zusätzlichen FUs werden bei phylogenetischen Vergleichen der FU-f zugeordnet (siehe Abb. 30). Es ist davon auszugehen, dass Genduplikationsereignisse der „eigentlichen“, an der Bildung der Zylinderwand beteiligten FU-f für die Entstehung der sechs zusätzlichen funktionellen Einheiten verantwortlich sind. Dieses Szenario wird an einer späteren Stelle dieses Kapitels (siehe: 4.5) ausführlicher diskutiert.

#### 4.1 „Codon-Usage“ Analyse der MtH-Untereinheiten

Im Rahmen der Arbeiten von Lieb *et al.* (2010) wurde die Hämolymphe von 22 verschiedenen *Melanoides tuberculata* Individuen elektronenmikroskopisch sowie proteinbiochemisch untersucht. Dabei wurde festgestellt, dass sich die Mengenverhältnisse der beiden MtH-Untereinheiten (MtH<sub>400</sub> und MtH<sub>550</sub>) in der Hämolymphe der untersuchten Individuen deutlich unterschieden. Die Spanne reichte vom ausschließlichen Vorkommen der MtH<sub>550</sub>-Untereinheit bis zum nahezu ausgeglichenen Mengenverhältnis der beiden Untereinheiten. Deutlich sichtbar war jedoch immer ein Überschuss der MtH<sub>550</sub>-Untereinheit (Lieb *et al.*, 2010). Diese Beobachtungen spiegelten sich auch in den elektronenmikroskopischen Aufnahmen der Hämolymphe der untersuchten Individuen wieder. Es wurden variable Anteile der verschiedenen Assemblierungsstufen des Mega-Hämocyanins festgestellt. Weiterhin konnten sogar Individuen identifiziert werden, die ausschließlich Mega-Dekamere in ihrer Hämolymphe besaßen (Lieb *et al.*, 2010). Eine mögliche Ursache für die differentielle Expression der beiden Untereinheiten des *Melanoides tuberculata* Hämocyanins könnte die Verwendung unterschiedlicher Codons in den MtH<sub>400</sub>- und MtH<sub>550</sub>-Untereinheiten sein. Die Tatsache, dass die Verwendung bestimmter Codons die Translationseffizienz beeinflussen kann, ist schon seit längerem bekannt und wurde sowohl für Prokaryoten (Robinson *et al.*, 1984) als auch für Eukaryoten (Neafsey *et al.*, 2007) untersucht und nachgewiesen. Daher wurden die beiden im Rahmen dieser Arbeit sequenzierten cDNAs der MtH-Untereinheiten statistisch auf die Häufigkeit der für die verschiedenen Aminosäuren kodierenden Codons untersucht. Darüber hinaus wurde die relative Aminosäurezusammensetzung der einzelnen FUs und der gesamten Untereinheiten statistisch ermittelt (siehe: 3.5 und Anhang I). Sowohl die

„Codon-Usage“-Analyse als auch die Analyse der Aminosäurezusammensetzung ließen jedoch keine auffälligen Unterschiede zwischen den beiden Untereinheiten des *Melanoides tuberculata* Hämocyanins erkennen.

**Tab.3: „Codon-Usage“-Analyse der vollständigen cDNA der MtH<sub>400</sub>- und MtH<sub>550</sub>-Untereinheiten**

Die Tabelle zeigt die für die 20 proteinogenen Aminosäuren kodierenden Codons sowie deren prozentuale Häufigkeit in den beiden cDNAs der MtH<sub>400</sub>- und MtH<sub>550</sub>-Untereinheit.

| Aminosäure         | Codon | MtH400 | MtH550 | Aminosäure    | Codon       | MtH400 | MtH550 |
|--------------------|-------|--------|--------|---------------|-------------|--------|--------|
| Alanin (A)         | GCG   | 7      | 7      | Asparagin (N) | AAT         | 13     | 11     |
|                    | GCA   | 5      | 4      |               | AAC         | 87     | 89     |
|                    | GCT   | 32     | 36     | Prolin (P)    | CCG         | 12     | 10     |
|                    | GCC   | 56     | 52     |               | CCA         | 9      | 6      |
| Cystein (C)        | TGT   | 19     | 28     |               | CCT         | 23     | 36     |
|                    | TGC   | 81     | 72     | CCC           | 56          | 47     |        |
| Asparaginsäure (D) | GAT   | 14     | 17     | Glutamin (Q)  | CAG         | 100    | 95     |
|                    | GAC   | 86     | 83     |               | CAA         | 0      | 5      |
| Glutaminsäure (E)  | GAG   | 85     | 84     | Arginin (R)   | AGG         | 19     | 15     |
|                    | GAA   | 15     | 16     |               | AGA         | 19     | 21     |
| Phenylalanin (F)   | TTT   | 16     | 15     |               | CGG         | 0      | 1      |
|                    | TTC   | 84     | 85     |               | CGA         | 4      | 4      |
| Glycin (G)         | GGG   | 4      | 1      |               | CGT         | 16     | 23     |
|                    | GGA   | 26     | 21     | CGC           | 42          | 36     |        |
|                    | GGT   | 20     | 24     | Serin (S)     | AGT         | 4      | 6      |
|                    | GGC   | 51     | 54     |               | AGC         | 33     | 26     |
| Histidin (H)       | CAT   | 7      | 13     |               | TCG         | 4      | 9      |
|                    | CAC   | 93     | 87     |               | TCA         | 2      | 4      |
| Isoleucin (I)      | ATA   | 2      | 0      |               | TCT         | 14     | 18     |
|                    | ATT   | 13     | 11     | TCC           | 42          | 37     |        |
|                    | ATC   | 85     | 89     | Threonin (T)  | ACG         | 21     | 19     |
| Lysin (K)          | AAG   | 89     | 89     |               | ACA         | 6      | 3      |
|                    | AAA   | 11     | 11     |               | ACT         | 21     | 22     |
| Leucin (L)         | TTG   | 3      | 5      |               | ACC         | 51     | 56     |
|                    | TTA   | 0      | 0      | Valin (V)     | GTG         | 41     | 40     |
|                    | CTG   | 59     | 59     |               | GTA         | 2      | 3      |
|                    | CTA   | 1      | 2      |               | GTT         | 17     | 17     |
|                    | CTT   | 10     | 10     |               | GTC         | 41     | 40     |
|                    | CTC   | 28     | 25     | Tryptophan(W) | TGG         | 100    | 100    |
| Methionin (M)      | ATG   | 100    | 100    |               | Tyrosin (Y) | TAT    | 14     |
|                    |       |        |        | TAC           |             | 86     | 91     |

Auch der Vergleich der Aminosäurezusammensetzung der beiden MtH-Untereinheiten und der Untereinheiten der Hämocyanine von *Aplysia californica*, *Haliotis tuberculata*, *Megathura crenulata* und *Octopus dofleini* zeigte keine deutlichen Unterschiede. Diese Daten sprechen dafür, dass die Verwendung bestimmter Codons bzw. eine unterschiedliche Aminosäurezusammensetzung als Ursache für die differentielle Expression der beiden MtH-Untereinheiten keine Rolle spielt. Vielmehr scheint die unterschiedliche Expression der beiden Untereinheiten des MtH in den untersuchten Individuen auf der Ebene der Transkription oder der mRNA-Prozessierung reguliert zu sein. Gerade die Initiation der Transkription ist ein entscheidender Schritt in der Kontrolle der Genexpression (Lemon *et al.*, 2000; Hochheimer *et al.*, 2003) und könnte auch im Falle der variablen Expression der MtH-Untereinheiten eine Rolle spielen.

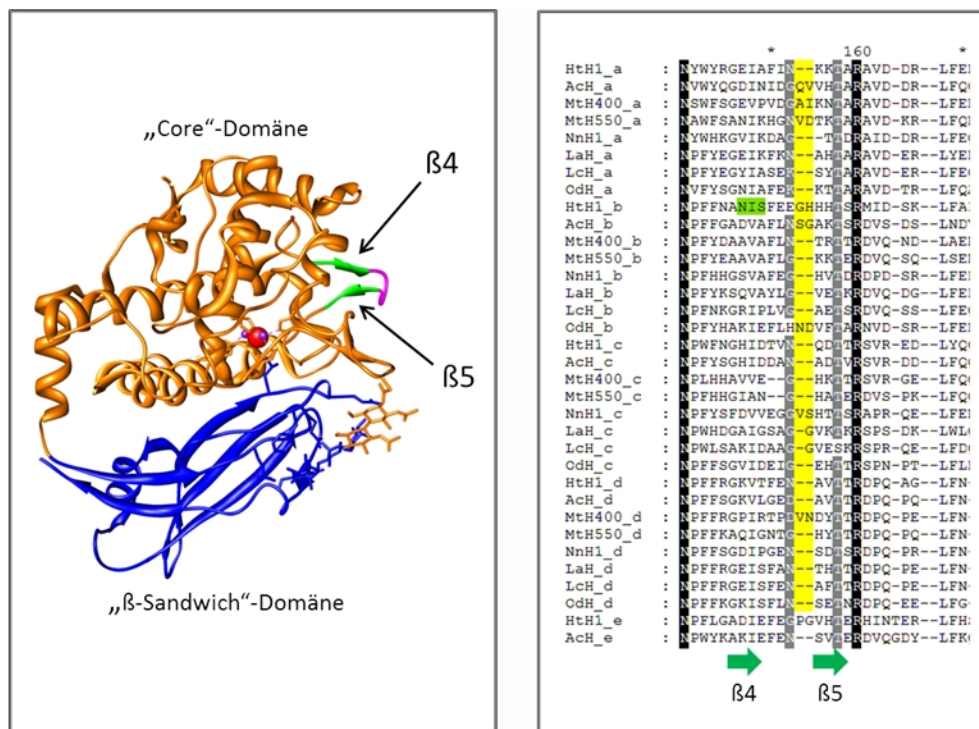
Ein nächster, wichtiger Schritt zur Aufklärung der individuell unterschiedlichen Expression der Untereinheiten des *Melanoides tuberculata* Hämocyanins ist die Identifizierung des Biosyntheseortes des Mega-Hämocyanins. Auch die Untersuchung der gewebsspezifischen Expression mittels „*In-situ*“-Hybridisierungsexperimenten könnte für die Klärung dieser Fragestellung hilfreich sein.

## 4.2 Vergleichende Sequenzanalyse des MtH

Die im Rahmen dieser Arbeit ermittelten cDNA-Sequenzen der FUs des MtH wurden in die abgeleiteten Aminosäuresequenzen übersetzt und mit den Sequenzen der FUs von *Haliotis tuberculata* (HtH), *Aplysia californica* (AcH), *Nucula nucleus* (NnH), *Leptochiton assellus* (LaH), *Lepitochitona cinerea* (LcH) sowie *Octopus dofleini* (OdH) aligniert. Besonderes Augenmerk wurde dabei auf das Vorkommen von charakteristischen Insertionen bzw. Deletionen („Indels“) in der Sequenz des MtH gerichtet. Im Folgenden werden die Ergebnisse dieser Analyse diskutiert. Das vollständige Sequenzalignment ist dem Anhang dieser Arbeit zu entnehmen (H).

Die Core-Domäne der funktionellen Einheiten weist einen exponierten Loop auf, welcher aus den kurzen  $\beta$ -Faltblättern „4“ und „5“ sowie einem dazwischenliegenden Sequenzabschnitt gebildet wird (Abb. 43). Dieser Loop wurde bereits von Cuff *et al.* (1998) als Ort mit einem häufigen Vorkommen von Insertionen und Deletionen beschrieben. Beide FU-a des MtH verfügen ebenso wie die FU-d des MtH<sub>400</sub> an Alignmentposition 153 bis 154 über eine Insertion von zwei Aminosäuren Länge. Alle

weiteren FUs des Mth weisen an dieser Stelle keine Insertion auf. Beide FU-c des Mth weisen im Bereich dieses Loops sogar eine Deletion von jeweils zwei Aminosäuren auf.

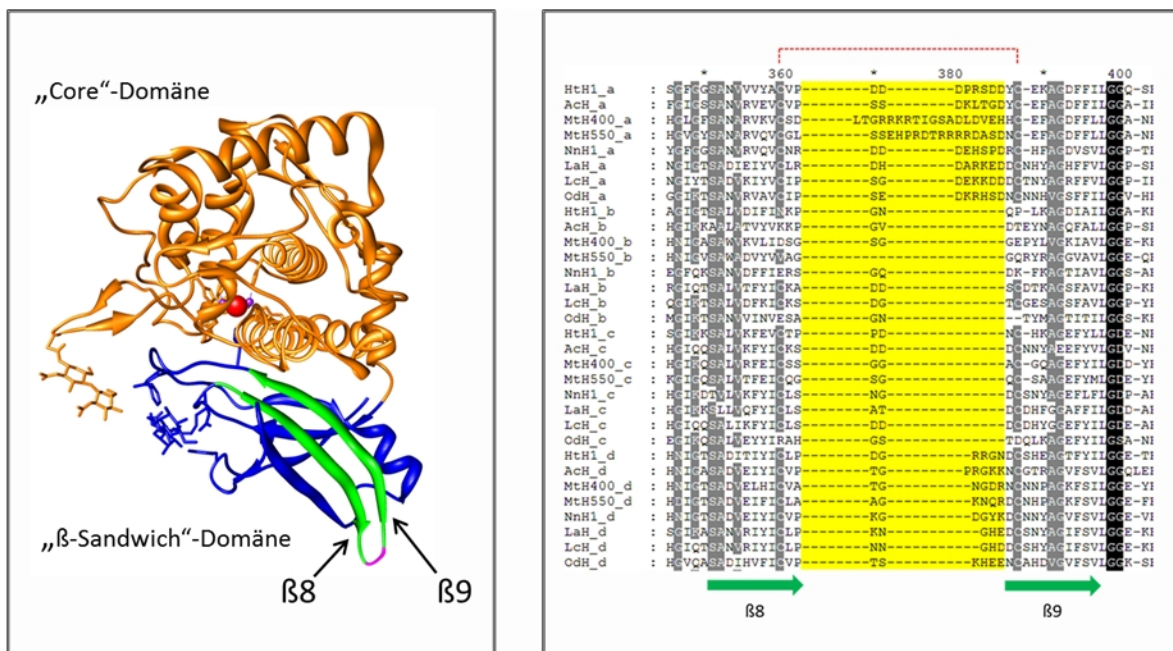


**Abb. 43: Indel-Position zwischen  $\beta 4$  und  $\beta 5$  sowie der entsprechende Ausschnitt aus dem Sequenzalignment**

Die linke Seite der Abbildung zeigt die Röntgenstruktur einer funktionellen Einheit der Mollusken-Hämocyanine (nach Cuff *et al.*, 1998; PDB-ID: 1JS8). Die Core-Domäne (orange) weist einen exponierten Loop (magenta) zwischen den beiden  $\beta$ -Faltblättern „4“ und „5“ (grün) auf. Die  $\beta$ -Sandwich-Domäne ist blau dargestellt. Auf der rechten Seite der Abbildung ist der korrespondierende Ausschnitt aus dem Sequenzalignment verschiedener Hämocyanin-FUs dargestellt. Die Positionen der  $\beta$ -Faltblätter „4“ und „5“ sind durch grüne Pfeile gekennzeichnet, während die Lage der Indels gelb markiert ist. Potentielle N-Glykosylierungsstellen (siehe: 4.3) sind im Alignment grün hinterlegt. Die verwendeten Abkürzungen der Hämocyanine werden im Anhang erläutert.

Eine weitere, wesentlich markantere Insertion befindet sich bei beiden FU-a des Mth zwischen den  $\beta$ -Faltblättern „8“ und „9“ der  $\beta$ -Sandwich-Domäne (Abb. 44). Diese Stelle bildet das „Anker-Interface“ zwischen den FU-a und FU-d und dient der Verknüpfung von zwei benachbarten Untereinheiten-Dimeren. Im Falle des KLH1 sowie des NpH wurde dieses wichtige Interface der Zylinderwand eingehend analysiert und beschrieben (Gatsogiannis *et al.*, 2007; Meissner *et al.*, 2007; Gatsogiannis, & Markl 2009). Das Alignment (Abb. 44) zeigt deutlich die bei den FU-a und FU-d auftretende Verlängerung des Loops. Die FU-a weist bei den bisher untersuchten Hämocyaninen sechs zusätzliche Aminosäuren auf, während die FU-d an dieser Stelle bis zu fünf zusätzliche Aminosäuren beinhaltet. Deutlich zu erkennen ist der massiv vergrößerte Loop der beiden FU-a des Mth

(Abb. 44). Die FU-a der 400 kDa Untereinheit weist 16 zusätzliche Aminosäuren auf, während die FU-a der 550 kDa Untereinheit über 14 zusätzliche Aminosäuren verfügt.

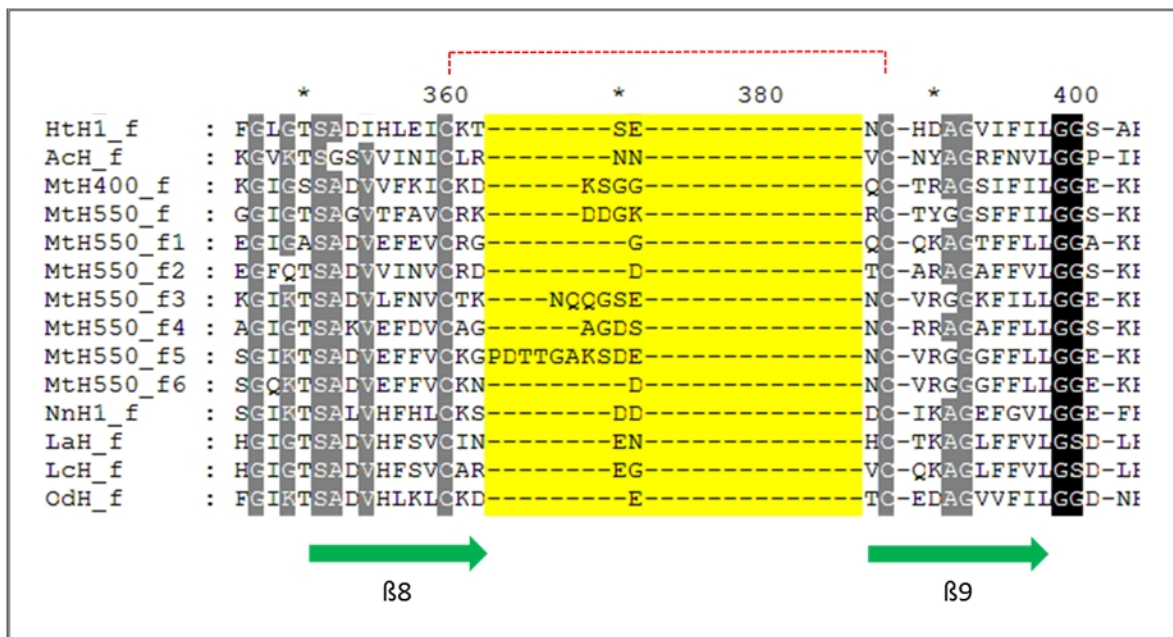


**Abb. 44: Indel-Position zwischen  $\beta 8$  und  $\beta 9$  sowie der entsprechende Ausschnitt aus dem Sequenzalignment**

Die linke Seite der Abbildung zeigt die Röntgenstruktur einer funktionellen Einheit der Mollusken-Hämocyanine (nach Cuff *et al.*, 1998; PDB-ID: 1JS8). Die  $\beta$ -Sandwich-Domäne (blau) weist einen exponierten Loop (magenta) zwischen den beiden  $\beta$ -Faltblättern „8“ und „9“ (grün) auf. Bei diesem Loop handelt es sich um das „Anker-Interface“ (siehe Haupttext). Die Core-Domäne ist orange dargestellt. Auf der rechten Seite der Abbildung ist der korrespondierende Ausschnitt aus dem Sequenzalignment verschiedener Hämocyanin-FUs dargestellt. Die Positionen der  $\beta$ -Faltblätter „8“ und „9“ sind durch grüne Pfeile gekennzeichnet, während die Lage der Indels gelb markiert ist. Zwischen den beiden Cysteinresten an den Alignmentpositionen 359 und 387 existiert eine Disulfidbrücke, welche in der Abbildung (rechts) durch eine unterbrochene rote Linie dargestellt ist. Die verwendeten Abkürzungen der Hämocyanine werden im Anhang erläutert.

Auffällig ist auch ein in beiden Loops der FU-a des MtH vorhandener Cluster von basischen Aminosäuren. Ein derartiges Motiv, bei dem sich jeweils vier basische Aminosäuren aneinander lagern („Arg-Arg-Lys-Arg“ bzw. „Arg-Arg-Arg-Arg“) wurde bisher bei keinem anderen Hämocyanin an dieser Position beobachtet. Die korrespondierende Stelle im entsprechenden Loop der FU-d des MtH weist bei beiden Untereinheiten vorwiegend polare bzw. neutrale Aminosäuren auf. Somit ergeben sich auch beim MtH deutliche Bindungsmöglichkeiten zwischen den beiden Loops. Auch verschiedene FU-f des MtH weisen an der beschriebenen Stellen zwischen den  $\beta$ -Faltblättern „8“ und „9“ ein markantes Indel auf (Abb. 45). Die Länge des Indels ist bei den verschiedenen FUs sehr variabel. Bei den FU-f<sub>1</sub>, FU-f<sub>2</sub> und FU-f<sub>6</sub> beträgt die Länge des Loops genauso wie bei den anderen in das Alignment einbezogenen Arten bis zu zwei

Aminosäuren. Die beiden eigentlichen FU-f der Mth Untereinheiten weisen genauso wie die FU-f<sub>4</sub> einen Loop von vier Aminosäuren Länge auf. Im Fall der FU-f<sub>3</sub> und FU-f<sub>6</sub> ist dieser Bereich mit sechs bzw. zehn Aminosäuren noch einmal deutlich länger. Wahrscheinlich führen die markanten Insertionen in der FU-f<sub>3</sub> und FU-f<sub>6</sub> zu einer erhöhten Flexibilität der entsprechenden FUs. Dies könnte für die korrekte Positionierung der FUs innerhalb des deutlich vergrößerten Kragenkomplexes des Mega-Dekamers erforderlich sein und würde damit auch dessen Stabilität erhöhen.

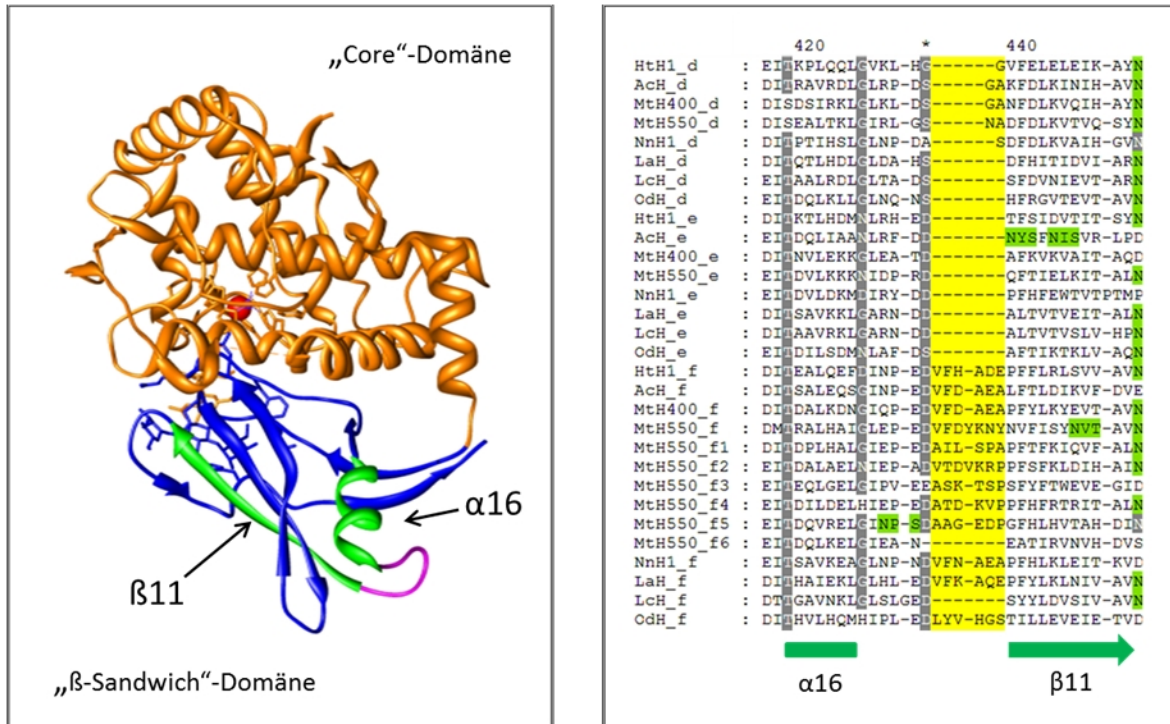


**Abb. 45: Ausschnitt aus dem Bereich der FU-f des Sequenzalignments**

Die Abbildung zeigt einen Ausschnitt aus dem Sequenzalignment verschiedener Hämocyanin-FU-f mit einem markanten Indel der Mth FUs. Die Positionen der  $\beta$ -Faltblätter „8“ und „9“ sind durch grüne Pfeile gekennzeichnet, während die Lage der Indels gelb markiert ist. Zwischen den beiden Cysteinresten an den Alignmentpositionen 359 und 387 existiert eine Disulfidbrücke, welche in der Abbildung durch eine unterbrochene rote Linie dargestellt ist. Die verwendeten Abkürzungen der Hämocyanine werden im Anhang erläutert.

Die beiden FU-d des Mth weisen eine zwei Aminosäuren umfassende Insertion im Bereich des Loops zwischen den Sekundärstrukturelementen  $\alpha$ 16 und  $\beta$ 11 auf (Abb. 46). Im Vergleich mit den in das Alignment einbezogenen Arten wurde dieses Indel auch bei *Aplysia californica* gefunden. *Nucula nucleus* und *Haliotis tuberculata* weisen in dem entsprechenden Bereich eine Indel von einer Aminosäure Länge auf. An exakt dieser Stelle in der  $\beta$ -Sandwich-Domäne befindet sich auch ein Indel in den meisten funktionellen Einheiten „f“ (Abb. 46). Im Fall des Mth verfügen die FU-f der 400 kDa Untereinheit sowie die FU-f<sub>1</sub>, FU-f<sub>3</sub>, FU-f<sub>4</sub> und FU-f<sub>5</sub> der 550 kDa Untereinheit über eine sechs Aminosäuren umfassende Insertion. Bei der FU-f und FU-f<sub>2</sub> der Mth<sub>550</sub>-Untereinheit hat

die Insertion eine Länge von sieben Aminosäuren. Die FU-f<sub>6</sub> weist an dieser Stelle keine Insertion auf. Bis auf *Lepitochitona cinerea* verfügen auch alle in das Alignment einbezogenen Arten zwischen der  $\alpha$ -Helix 16 und dem  $\beta$ -Faltblatt 11 über eine Insertion von sechs Aminosäuren Länge.



**Abb. 46: Indel-Position zwischen  $\alpha$ 16 und  $\beta$ 11 sowie der entsprechende Ausschnitt aus dem Sequenzalignment**

Die linke Seite der Abbildung zeigt die Röntgenstruktur einer funktionellen Einheit der Mollusken-Hämocyanine (nach Cuff *et al.*, 1998; PDB-ID: 1JS8). Die  $\beta$ -Sandwich-Domäne (blau) weist einen exponierten Loop (magenta) zwischen der  $\alpha$ -Helix „16“ und dem  $\beta$ -Faltblatt „11“ (grün) auf. Die Core-Domäne ist orange dargestellt. Auf der rechten Seite der Abbildung ist der korrespondierende Ausschnitt aus dem Sequenzalignment verschiedener Hämocyanin-FUs dargestellt. Die Positionen der oben genannten Sekundärstrukturelemente sind durch grüne Symbole (Rechteck =  $\alpha$ -Helix, Pfeil =  $\beta$ -Faltblatt) gekennzeichnet, während die Lage der Indels gelb markiert ist. Potentielle N-Glykosylierungsstellen (siehe: 4.3) sind im Alignment grün hinterlegt. Die verwendeten Abkürzungen der Hämocyanine werden im Anhang erläutert.

Das Indel der FU-f wurde bereits in den Arbeiten von Möller (2008) und Streit (2008) beschrieben und diskutiert. Bei ihren Untersuchungen zur molekularen Phylogenie der Polyplacophora diskutierte Möller (2008) einen Zusammenhang zwischen dem Besitz dieser Insertion und der Fähigkeit Didekamere zu bilden. Da das Hämocyanin von *Octopus dofleini* die FU-a bis FU-g umfasst und die beschriebene Insertion aufweist aber dennoch keine Didekamere ausbildet, scheint auch das Vorhandensein der FU-h essentiell für den Prozess der Didekamerbildung zu sein (Möller, 2008).

Bisher ist der Verlauf der FU-f<sub>1</sub> bis FU-f<sub>6</sub> in der Quartärstruktur des Mega-Dekamers nicht bekannt. Daher kann an dieser Stelle noch keine Aussage über die Lage der entsprechenden, die Insertion beinhaltenden Loops im Mega-Hämocyanin getroffen werden. Es ist jedoch denkbar, dass diese Loops eine exponierte Position im Mega-Dekamer einnehmen und für die Assemblierung und Stabilisierung der Mega-Tridekamere mit verantwortlich sind.

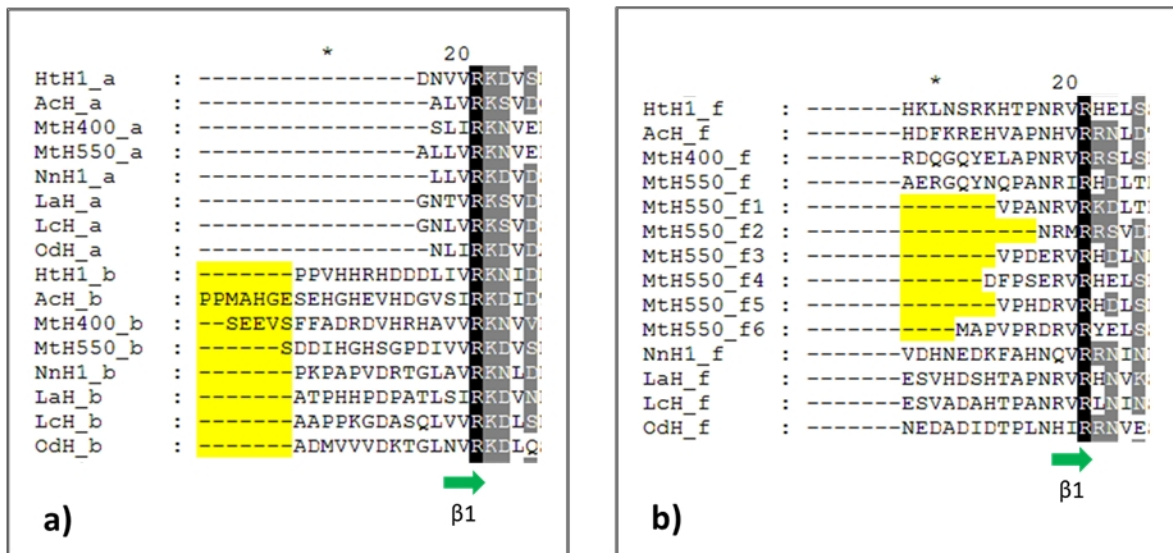
|           |   |         |   |   |
|-----------|---|---------|---|---|
| HtH1_f    | : | VFC-ACE | : | 6 |
| AcH_f     | : | VFC-AEA | : | 6 |
| MtH400_f  | : | VFC-AEA | : | 6 |
| MtH550_f  | : | VFYKNI  | : | 7 |
| MtH550_f1 | : | AIL-SPA | : | 6 |
| MtH550_f2 | : | VTVKRP  | : | 7 |
| MtH550_f3 | : | ASK-TSP | : | 6 |
| MtH550_f4 | : | ATL-KVP | : | 6 |
| MtH550_f5 | : | AAG-EFP | : | 6 |
| MtH550_f6 | : | -----   | : | - |
| NnH1_f    | : | VFN-AEA | : | 6 |
| LaH_f     | : | VFK-ACE | : | 6 |
| LcH_f     | : | -----   | : | - |
| OdH_f     | : | LVV-HGS | : | 6 |

**Abb. 47: Physikalisch-chemische Eigenschaften der Aminosäuren des Indels zwischen  $\alpha 16$  und  $\beta 11$**

Die Abbildung zeigt die alignierten Sequenzen der zwischen den Strukturelementen  $\alpha 16$  und  $\beta 11$  liegenden Indels verschiedener FU-f. Die Aminosäuren sind entsprechend ihrer physikalisch-chemischen Eigenschaften farblich markiert (unpolar/hydrophob = gelb, polar/neutral = grün, basisch = blau, sauer = rot). Den funktionellen Einheiten MtH<sub>550</sub>-f<sub>6</sub> und LcH-f fehlt das entsprechende Indel. Deutlich erkennbar ist der große Anteil hydrophober Aminosäuren. Die verwendeten Abkürzungen der Hämocyanine werden im Anhang erläutert.

Es ist denkbar, dass die FU-f<sub>6</sub>, die an der erwähnten Stelle keine Insertion aufweist, eine terminale Position in der Untereinheit einnimmt und daher nicht an der Assemblierung und/oder Stabilisierung der Mega-Tridekamere beteiligt ist.

Um Hinweise auf die funktionelle Bedeutung der beschriebenen Insertion der FU-f zu erhalten, wurden die entsprechenden Sequenzen auf ihre Aminosäurezusammensetzung untersucht (Abb. 48). Deutlich erkennbar ist bei den meisten Sequenzen ein klarer Überschuss hydrophober Aminosäuren (Anteil: 29 % bis 83 %) der auch von Möller (2008) bei molekularphylogenetischen Arbeiten an Polyplacophoren beobachtet werden konnte. Alle Indels verfügen zudem über einzelne saure bzw. basische Aminosäuren (Abb. 47; rote bzw. blaue Kennzeichnung). Diese könnten ähnlich wie im Fall des KLH1 (Gatsogiannis & Markl, 2009) durch die Bindung von divalenten Kationen bzw. durch die Ausbildung von Salzbrücken an der Bildung der Di- und Tridekamere beteiligt sein.



**Abb. 48: Ausschnitt aus dem Bereich der Peptidlinker des Sequenzalignments**

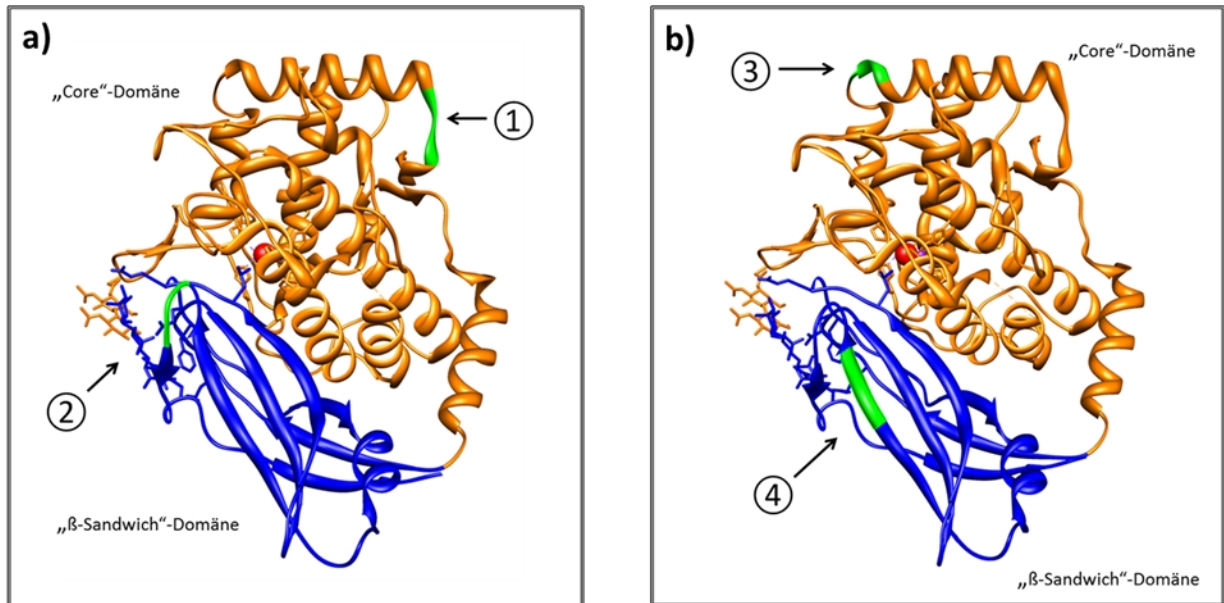
Die Abbildung zeigt den Bereich der Peptidlinker der FU-a, FU-b (**a**) und FU-f (**b**) aus dem Sequenzalignment verschiedener Hämocyanin-FUs. Die Position des  $\beta$ -Faltblattes „1“ ist durch einen grünen Pfeil gekennzeichnet, während die Lage der Indels gelb markiert ist. Die verwendeten Abkürzungen der Hämocyanine werden im Anhang erläutert.

Auch im Bereich der Peptidlinker, die die einzelnen FUs miteinander verbinden, weist das Mth einige auffällige Indels auf (Abb. 48). Die beiden FU-b der Mth-Untereinheiten sind im Vergleich zu den in das Alignment einbezogenen Spezies am N-Terminus um bis zu fünf Aminosäuren verlängert. Eine solche Verlängerung der FU-b kann außer bei den Mth-Sequenzen nur noch bei *Aplysia californica* beobachtet werden.

Auch die sechs zusätzlichen FU-f der 550 kDa Untereinheit verfügen über deutliche Indels im Bereich der Linker-Peptide. Die Länge der Deletionen ist bei den FUs unterschiedlich ausgeprägt und kann wie im Fall der FU-f<sub>2</sub> bis zu zehn Aminosäuren betragen. Die Verkürzung der Peptidlinker könnte zum Beispiel für eine kompaktere Anordnung der FUs innerhalb des vergrößerten Kragenkomplexes des Mega-Dekamers erforderlich sein und somit wesentlich zur Stabilität des Mega-Dekamers beitragen.

### 4.3 Analyse der potentiellen N-Glykosylierungsstellen

Die Mth<sub>400</sub>-Untereinheit besitzt sechs potentielle N-Glykosylierungsstellen, während die Mth<sub>550</sub>-Untereinheit über insgesamt zwölf dieser Motive verfügt (siehe: 3.4). Durch den Vergleich der FUs des *Melanoides tuberculata* Hämocyanins mit der Sequenz der FU-g des Hämocyanins von *Octopus dofleini* wurde die Lage der N-Glykosylierungsstellen des Mth innerhalb der 3D-Struktur der OdH-g (Cuff *et al.*, 1998) bestimmt.

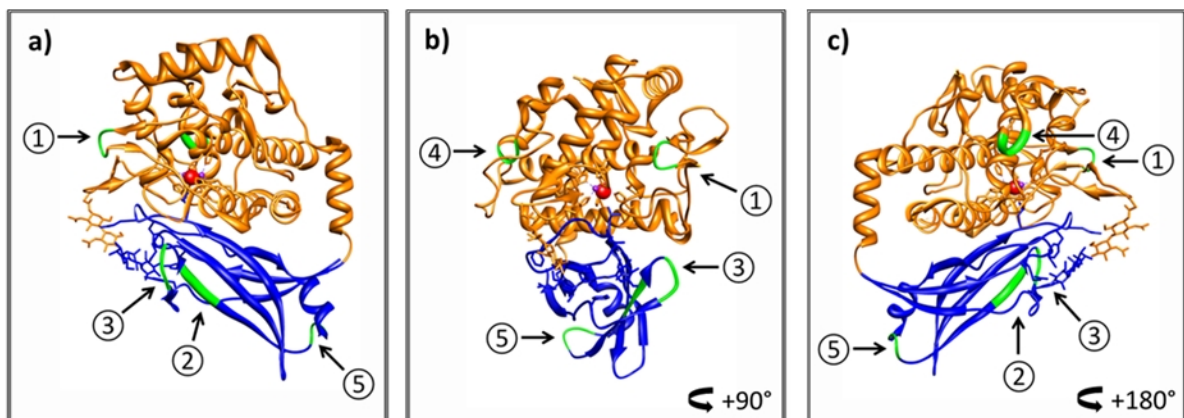


**Abb. 49: Schemata der OdH-g mit Projektion der N-Glykosylierungsstellen der FU-a und FU-b des Mth**

Die Grafik zeigt schematisch die Röntgenstruktur der FU-g des *Octopus dofleini* Hämocyanins (nach Cuff *et al.*, 1998; PDB-ID: 1JS8). Die vorwiegend  $\alpha$ -helikale „Core-Domäne“ ist orange, die „ $\beta$ -Sandwich-Domäne“ blau gekennzeichnet. Die potentiellen N-Glykosylierungsstellen der FU-a (a) und FU-b (b) des Mth sind grün dargestellt. Die Mth<sub>400</sub>-a verfügt über eine N-Glykosylierungsstelle zwischen Helix  $\alpha$ 1 und  $\alpha$ 2 (1) sowie zwischen den  $\beta$ -Faltblättern  $\beta$ 11 und  $\beta$ 12 (2). An letztgenannter Position befindet sich auch die einzige N-Glykosylierungsstelle der Mth<sub>550</sub>-a. Die FU-b beider Mth-Untereinheiten weisen eine N-Glykosylierungsstelle zwischen den Helices  $\alpha$ 2 und  $\alpha$ 3 auf (3). Zusätzlich verfügt die FU-b der Mth<sub>550</sub>-Untereinheit über ein entsprechendes Motiv innerhalb des  $\beta$ -Faltblattes  $\beta$ 11 (4).

Auffällig ist, dass sich elf der insgesamt achtzehn N-Glykosylierungsstellen (61 %) des Mth auf die exponierte Region zwischen den beiden  $\beta$ -Faltblättern  $\beta$ 11 und  $\beta$ 12 der  $\beta$ -Sandwich-Domäne konzentrieren (Abb. 49 (2)). Die FU-a, FU-d und FU-f beider Untereinheiten sowie die Mth<sub>400</sub>-g, Mth<sub>550</sub>-e, Mth<sub>550</sub>-f<sub>1</sub>, Mth<sub>550</sub>-f<sub>2</sub> und Mth<sub>550</sub>-f<sub>4</sub> weisen hier jeweils ein Motiv der Form NGS/T auf. Eine Häufung von N-Glykosylierungsstellen in diesem Bereich wurde bereits für die FU-a, FU-d, FU-e und FU-f des Hth1 (Lieb *et al.*, 2000) sowie die FU-a, FU-d, FU-e, FU-f und FU-g des Hth2 beschrieben (Altenhein *et al.*, 2002). Im Falle des Hämocyanins von *Aplysia californica* konnten für die FU-a, FU-d und

FU-g potentielle N-Glykosylierungsstellen an der entsprechenden Position identifiziert werden (Lieb *et al.*, 2004). Auch für die Hämocyanine weiterer Mollusken wurden N-Glykosylierungsstellen zwischen den  $\beta$ -Faltblättern  $\beta 11$  und  $\beta 12$  beschrieben (Miller *et al.*, 1998; Bergmann *et al.*, 2006; Streit *et al.*, 2003; Lieb *et al.*, 2004; Bergmann *et al.*, 2007; Gatsogiannis & Markl, 2009). Diese Stelle, bei der es sich um die am häufigsten mit potentiellen N-Glykosylierungsstellen besetzte Position bei Mollusken-Hämocyaninen handelt, scheint auch eine Rolle bei der Stabilisierung der Tertiärstruktur des Hämocyanins zu spielen (Lieb *et al.*, 2000; Bergmann *et al.*, 2007).



**Abb. 50: Schemata der OdH-g mit Projektion der N-Glykosylierungsstellen der FU-f des Mth**

Die Grafik zeigt schematisch die Röntgenstruktur der FU-g des *Octopus dofleini* Hämocyanins (nach Cuff *et al.*, 1998; PDB-ID: 1JS8). Die Ansichten sind im Bezug zu (a) um  $90^\circ$  (b) bzw.  $180^\circ$  (c) um die y-Achse gedreht. Die vorwiegend  $\alpha$ -helikale „Core-Domäne“ ist orange, die „ $\beta$ -Sandwich-Domäne“ blau gekennzeichnet. Die potentiellen N-Glykosylierungsstellen der FU-f des Mth sind grün dargestellt. Die Mth<sub>400</sub>-f sowie die Mth<sub>550</sub>-f verfügen über eine N-Glykosylierungsstelle zwischen den  $\beta$ -Faltblättern  $\beta 11$  und  $\beta 12$  (③). An dieser Position befinden sich auch die einzigen N-Glykosylierungsstellen der Mth<sub>550</sub>-f<sub>1</sub>, -f<sub>2</sub> und -f<sub>4</sub>. Die Mth<sub>550</sub>-f weist zusätzlich eine potentielle N-Glykosylierungsstelle im Bereich des  $\beta$ -Faltblattes  $\beta 11$  auf (②). Zwischen den  $\beta$ -Faltblättern  $\beta 4$  und  $\beta 5$  liegt die einzige N-Glykosylierungsstelle der C-terminalen FU-f<sub>5</sub> und FU-f<sub>6</sub> (①). Motive der Art „NPS“, die aufgrund des Prolins die N-Glykosylierung vollständig zu blockieren scheinen (Bause, 1983) sind sowohl zwischen den  $\alpha$ -Helices  $\alpha 12$  und  $\alpha' 13$  (④); Mth<sub>400</sub>-f) als auch zwischen den Sekundärstrukturelementen  $\alpha 16$  und  $\beta 11$  (⑤); Mth<sub>550</sub>-f<sub>5</sub>) lokalisiert.

Beide FU-b des Hämocyanins von *Melanoides tuberculata* besitzen eine potentielle N-Glykosylierungsstelle im N-terminalen Bereich der „Core-Domäne“ zwischen den  $\alpha$ -Helices  $\alpha 2$  und  $\alpha' 3$  (Abb. 51 (③)). Auch die FU-f des Hth1 weist an exakt dieser exponierten Position eine N-Glykosylierungsstelle auf (Lieb *et al.*, 2000). Zusätzlich verfügt die Mth<sub>550</sub>-b genauso wie die Mth<sub>550</sub>-f über eine N-Glykosylierungsstelle im Bereich des  $\beta$ -Faltblattes  $\beta 11$  (Abb. 49 (④); Abb. 50 (②)). Auch in diesem Falle liegt der entsprechende Bereich leicht zugänglich auf der Außenseite der funktionellen Einheit. Eine weitere exponierte N-Glykosylierungsstelle, die zwischen den  $\alpha$ -Helices  $\alpha' 1$  und  $\alpha 2$  am

N-terminalen Ende der „Core-Domäne“ liegt, findet sich in der FU-a der M<sub>tH</sub><sub>400</sub>-Untereinheit (Abb. 49 ①). Die FU-a der 550 kDa Untereinheit weist ebenso wie die funktionellen Einheiten M<sub>tH</sub><sub>400</sub>-e und -h, im Gegensatz zu vielen bisher untersuchten Gastropoden-Hämocyaninen, keine potentiellen N-Glykosylierungsstellen auf (Lieb *et al.*, 2000; Altenhein *et al.*, 2002; Lieb *et al.*, 2004; Gatsogiannis & Markl, 2009). Auch beide FU-c des M<sub>tH</sub> sowie die M<sub>tH</sub><sub>550</sub>-f<sub>3</sub> besitzen keine N-Glykosylierungsstellen. Das Fehlen von potentiellen N-Glykosylierungsstellen in den FU-c scheint typisch für Mollusken-Hämocyanine zu sein (Lieb *et al.*, 2004). Im Falle der KLH2-c scheinen stattdessen O-glykosidisch gebundene Zucker vorhanden zu sein (Stoeva *et al.*, 1999). Ob dies auch auf das Hämocyanin von *Melanoides tuberculata* zutrifft muss in weiterführenden Untersuchungen geklärt werden. Die C-terminalen FU-f<sub>5</sub> und FU-f<sub>6</sub> der M<sub>tH</sub><sub>550</sub>-Untereinheit verfügen über jeweils eine potentielle N-Glykosylierungsstelle an der im Vergleich zu anderen Mollusken-Hämocyaninen höchst ungewöhnlichen Position zwischen den  $\beta$ -Faltblättern  $\beta_4$  und  $\beta_5$  (Abb. 50 ①). Dieser Bereich der „Core-Domäne“ bildet einen exponierten Loop, der deutlich aus der Röntgenstruktur heraussteht. Es ist denkbar, dass die FU-f<sub>5</sub> und FU-f<sub>6</sub> aufgrund ihrer C-terminalen Lage in der Untereinheit an der distalen Spitze des vergrößerten Kragenkomplexes des Mega-Dekamers positioniert sind und sich daher die potentiellen N-Glykosylierungen in einem von außerhalb des Proteins gut zugänglichen Bereich befinden.

Außer den vorangehend beschriebenen potentiellen N-Glykosylierungsstellen verfügt das M<sub>tH</sub> über zwei „NPS“-Motive innerhalb der funktionellen Einheiten M<sub>tH</sub><sub>400</sub>-f und M<sub>tH</sub><sub>550</sub>-f<sub>5</sub>. Diese Sequenzmotive verfügen normalerweise nicht über N-glykosidisch gebundene Zucker und scheinen die N-Glykosylierung an den entsprechenden Stellen aufgrund des enthaltenen Prolins sogar zu blockieren (Bause, 1983).

#### 4.4 Analyse der Dekamer-Dekamer Interfaces

Mit Hilfe einer hochaufgelösten 3D-Rekonstruktion (9 Å) sowie eines molekularen Modells des KLH1-Didekamers konnten vor kurzem erstmalig die Kontaktstellen zwischen den beiden Hälften eines Hämocyanin-Didekamers genau beschrieben werden (Gatsogiannis & Markl, 2009). Die am Zylinderrand gelegenen FU-a, FU-b, FU-c und FU-f bilden dabei insgesamt 45 Kontaktstellen zwischen den beiden Dekameren aus, die sich in folgende vier verschiedene Interface-Typen einteilen lassen: „b<sub>2</sub>/a<sub>2</sub>“, „c<sub>2</sub>/a<sub>2</sub>“, „f<sub>2</sub>/c<sub>2</sub>“ und „f<sub>2</sub>/f<sub>2</sub>“ (Gatsogiannis & Markl, 2009).

Die 400 kDa Untereinheit des MtH ist in der Lage sowohl zusammen mit der 550 kDa Untereinheit Mega-Di- und Tridekamere als auch in isolierter Form typische Gastropoden-Hämocyanin Didekamere zu bilden. Daher wurden die FU-a, FU-b und FU-c sowie die beiden „eigentlichen“ FU-f mit den entsprechenden FUs des KLH1 aligniert und die an den Dekamer-Dekamer Kontaktstellen beteiligten Aminosäuren untersucht (Abb. 51).

Die Dimerisierung des KLH1 lässt sich über die  $\text{Ca}^{2+}$ - bzw.  $\text{Mg}^{2+}$ -Konzentration steuern (Swerdlow *et al.*, 1996; Harris *et al.*, 1997; Harris *et al.*, 1997). Für die Hämocyanine von *Nautilus pompilius* (NpH) und *Megathura crenulata* (KLH) konnten in der Zylinderwand verschiedene Carboxylat-Reste identifiziert werden, die durch die Brückenbildung mit einem divalenten Kation zur Stabilisierung der Quartärstruktur des Hämocyanins beitragen (Gatsogiannis *et al.*, 2007; Gatsogiannis & Markl, 2009). Dieser Mechanismus scheint auch bei den beiden Dekamer-Dekamer Interfaces „b<sub>2</sub>/a<sub>2</sub>“ und „f<sub>2</sub>/f<sub>2</sub>“ des KLH1 eine Rolle zu spielen (Gatsogiannis & Markl, 2009). Verantwortlich für die Bindung der Kationen sind bei diesen beiden Interfaces die Glutamate an den Alignmentpositionen 15 und 320 („b<sub>2</sub>/a<sub>2</sub>“) sowie 372 („f<sub>2</sub>/f<sub>2</sub>“) (Abb. 51). Das Glutamat an Position 15 ist in der MtH<sub>400</sub>-Untereinheit nicht vorhanden und in der MtH<sub>550</sub>-Untereinheit durch ein unpolares Alanin ersetzt, während das zweite beteiligte Glutamat (320) bei beiden Untereinheiten des MtH nachweisbar ist. Aufgrund der beobachteten Aminosäuresubstitutionen scheint diese Kontaktstelle in der für das KLH beschriebenen Form beim MtH nicht vorhanden zu sein. Das am Interface „f<sub>2</sub>/f<sub>2</sub>“ beteiligte Glutamat (372) wurde bei der MtH<sub>400</sub>-Untereinheit durch ein Glycin ersetzt und scheint somit kein Interface auszubilden. Die 550 kDa Untereinheit des MtH weist an dieser Position ein positiv geladenes Lysin auf, welches ebenso wie das Glycin der 400 kDa Untereinheit nicht der im Fall des KLH beschriebenen Kontaktstelle entspricht.

Das „f<sub>2</sub>/c<sub>2</sub>“ Interface beruht auf den ionischen Wechselwirkungen zwischen einem positiv geladenen Arginin (337) sowie einem negativ geladenen Glutamat (66). Bei beiden MtH-Untereinheiten ist das Glutamat durch ein Aspartat, welches die gleiche Ladung aufweist, ersetzt. Da jedoch statt des Arginins bei beiden MtH-Untereinheiten ein ungeladenes Glycin vorhanden ist, scheint auch diese Kontaktstelle beim MtH nicht an der Bildung und Stabilisierung der Di- bzw. Tridekamere beteiligt zu sein.

```

*          20          *          40          *          60          *          80          *
KLH1-a : -----ENLVRKSV EHLTQ EETL DLQAALRELQMDSSS IGFQKIAAA HCAPAS CVHKD --TSI ACCI HGMPTF PHWHRLAYVV
Mth400-a : -----SLIRKKNV ELLSS QEILSLQTS LRALAADE SASGFKA VAA YHCEPSM CEDNDG-VAI ACCI HGMVAVF PQWHRLYVV
Mth550-a : -----ALLVRKNV EDLTQ EEVLS LQSYLSKLEAD ESEYGFQKLAAY HCEPSM CQSDGD-TARACC I HGMANF PQWHRLYVV
KLH1-b : -----PPVKHHQS ANLLVVRKNI NDLTREEV LNREAF HFKQED RSV DGYQATAE YHGLPAR CPRPDAKDRY ACCV HGMPIF PHWHRLVVT
Mth400-b : SEEVSVFADR DVHRAVVRKNI VVLDTEEVYELRLAMD R FQNDI RSV DGYQAI AEYHGLPAR CPRPDAAVRY ACCI HGMATF PHWHRLFVV
Mth550-b : -----SDD IHGHS GPD IVVRKDV SELTAEEM YLRQAMSK FQNDI RSV DGYQAVAE F HGLPAR CPRPDAAVRY ACCI HGMATF PHWHRLFVV
KLH1-c : -----VKFDK VPRSRL IRKKNV DRLSPEEMNE LR KALALLKED K SAGGFQQLGAF HCEPKV CPSPEASKKF ACCV HGMVSVF PHWHRLLVV
Mth400-c : -----TFEVRGVSDK RTRLDVDHLNSE QIQNL RDAFKALEDD DTSINGFNQ IAAF HCEPSP CPSPEAETKY ACCI HGMPTF PHWHRLITV
Mth550-c : -----VHHDAGDAPS PDKVRSNI ESLTREQVQNL R DALQALEDD DSNAGFAQ IAAF HCEPRW CPSPEAEKKY ACCI HGMVAVF PHWHRLITV
KLH1-f : -----HDISSH HSLNKNVR HDLSTLSE RDIGSLKYALSSLQAD T SADGFAAIAS F HGLPAK CND S-HNNEV ACCI HGMPTF PHWHRLYLV
Mth400-f : -----RDQGGY ELAPNRVRRSLSDLTERD VMSLKSALHD LQEDD SATGWQSLAAY HGVPA LCPSP-EAKY ACCI HGMPTF PHWHRLYLV
Mth550-f : -----AERGQYNQ PANRIR HDLTHLSE RD FMSLKSALND LQAD TCKMGWQSLAS F HGVPA LCPPT-AEATY ACCI HGMPTF PHWHRLYLV
    
```

```

*          100         *          120        *          140         *          160         *          180
KLH1-a : HMERALQTKRRTSGLPYWDWTEPITQLPSLAADPVYIDSQ-GGKAHTN YWYRGN I--DFLDK--KTRAVDDRLFEKVKP GGHHTHLSV
Mth400-a : QFEQLLVNKG LNI GVPYWDWTEPLEPLTHLPSLVEEQ IFRDPN-GGVGKNSWF SGEV PVDGAIK--NTRAVDDRLFEKVA P GKN TQLFNII
Mth550-a : QFEQLLRKKG MMSMGV PYWDWTEGELTSLPDLVEEH IFRDPN-GGRKENAWF SANIKHGNDV T--KTRAVD KRLFQKVGDGEK TQLFNII
KLH1-b : QVEDALVGRGATIGIPYWDWTEPMTHI PGLAGNKTYVDSH--GASHTNPFHSSVI--AFEENAPHTKQIDQRLFKPATFGHHTDLFNQI
Mth400-b : QIEDELKAR DLHFGMPYWDWTPNAQI PAIAAEATYQDPHPHDVYTLHNP FYDAAV--AFLNT--RTTRD VQNDLAENPIYGDH SKLYDGI
Mth550-b : QIEDELKAR DLHIGV PYWDWTRPNTHI PKLAAEATYQDPHEGVT LNNPFYEA AV--AFLGK--KTRD VQVQSLSENPF GHD TDLFDGM
KLH1-c : QSENALR RHGYD GALPYWDWTSPLNHL PELADHEKYVDE-DGVEKHN PWF DGH I--DTVDK--TTTRSVQNKLF EQPEFGHYTSIAKV
Mth400-c : QLENALR RHGFE SGLPYWDWTEPMTELPE SFTAETYDNPV-SGANEANPLHHA VV--E--GH--KTRSVR GELFQ QPAFGK FTRI GEQV
Mth550-c : QAENALR RHGFEDGLPYWDWTRPMKAL PQVVE SAQYV NQQ-TGQ TMDNPFHGGIA--N--GH--ATERD VSPKLFQ QPEFGK FTDIAELV
KLH1-f : QFEQALRRH GSSVAV PYWDWTKPIHNI PHLFTDK EYD YW-RNKVMPNP FARGYV--PSHD T--YTRD VQVQGLFHLTSTGEHSALLNQA
Mth400-f : AVEHSLIKH GSSVAI PYWDWTEPLDKL PELFTKQTYD AW-RDEYV DNP FARGYV--KGDEA--YTRD PQPELKRSDGKHSV LFDQV
Mth550-f : LVEHALIEH GSSSAI PYWDWTEPLDGL PELFTKQTYD AW-KDKVFDNPF FARGYI--KSDVG--YTRD PQSQMSMRS DKGKHS I LFDVRV
    
```

```

*          200         *          220        *          240         *          260         *
KLH1-a : LDAL EQDFE KFEIQ FELAHNAIHYLVGGKHDYSMANLE YTA YDPIFFLHHSNVDRIFAIWQRLQELRNKDPKAMDCAQEL LHQMEPPFS
Mth400-a : LDAL EQDDY QFEVQ LEVAHNIHYLVGGRHQYSMATLEYSSYDPIFFLHHSNIDRIFVIVQELQKIRGKPKYDHSDCSVEQFRWNLKPPF
Mth550-a : LDAL EKDDY QFEVQ FEVAHNIHYLVGGRHKYSMATLEYSSYDPIFFLHHSAMDRI FVIWQELQKIRGKPSDHDACSVLEFRWNLKPPF
KLH1-b : LYAFEQEDY QDFEVQFEITHNTI HAWTGGSEHFSMSSLHYTAFDPLFFYHHSNVDRILWAVWQALQMRGKSPY-RAHCAI SLEHMHLPFFA
Mth400-b : LLAFEQTDY QDFEIQFEVIHNSPHFLVGGFGAYTLATLHYSAFDPIFFYHHSNVDRIMWAIWQELQIRRLGLPY-RAHCAHTMTHEPMKPPFS
Mth550-b : LLAFEQTDY QDFEIQFEVTHNAPHFLVGGFKPYSLATLHYSAFDPIFFYHHSNVDRILWAIWQELQIRRLGLPY-EAHCA NSMTHDLLKPPF
KLH1-c : LLALEQDNF QDFEIQYEIAHNYIHALVGGAGYLGMSALRYTAFDPLFFYHHSNVDRILWAIWQALQKIRGKPYNTANCAITQLRKLQPFF
Mth400-c : MLAPEQTDY QSFEVQFEVAHNYIHALVGGNELYSMASLRYTAYDPLFALHHSNVDRILWAIWQALQKIRGLPYNTANCAITQLRKLQPFFA
Mth550-c : MFAPEQTDY QDFEVQMEIAHNYIHALVGGSKPYSMSSLRYTAYDPIFFLHHSNVDRILWAIWQALQKIRGLPYNTANCAI LASMRLKQPFFA
KLH1-f : LLALEQHDY QDFAVQFEVMHNTIHYLVGGPQVYSLSSLHYASYDPIFFIHH SFV DKLWAVWQALQKIRGKPSERADCAVNYMYDDMHPFF
Mth400-f : LLALEQTDY QDFEVQYEVTHNAIHYLVGGPQTYSLSLHYSSYDPIFFVHHSFV DKLWAVWQALQKIRGKPSERADCAVNYMYDDMHPFF
Mth550-f : LDAL EQDDF QDFEVQFEVTHNAIHYLVGGPQYSLSSLHYSSYDPIFFVHHSFV DKLWAVWQELQRRR GKP YDRADCAVNYMSKPMMPFFN
    
```

```

*          280         *          300         *          320         *          340         *          360
KLH1-a : WED--NDIPLTNEHSTPADLFDY-CELHYDYD TLNLGNTPEELKTYLDERSSRAAFASFR LKGF GGSANV FVYVCI----PDDND--
Mth400-a : RDT--NPIALTKQHTAKELFSN-LQLGYVYD TLNLGMDVQLYNLMEQRRAREAFGVFSLHG CFSANARVVKCI----DLTGRRK
Mth550-a : RSS--NPVELTKQHSKADSVWHH-PELGYWYD TLTLGGMDLQLHAYLEKRRAREAYATFSLHGVGYSANARVQVCGLSSEHPRDTRRR
KLH1-b : FSSPLNNEKTHANAMPNKIYD YENVLHYTYEDLTFGGI SLENI EKMIHENQOEDRIYAGFLLAGIRTSANVDIFIK
Mth400-b : FAAPLNNEKTFSSHESQVYD YERDLGYD SLTFGGMTIEQENLEYERQSKERV FVGLHNI GASAWVKVLIDS-----
Mth550-b : FEAPLNNEKTHSHARP TKIYD YELELDYTYDNLQFGG MTIEQLEHYLEERQAREV FVGLFHNIGVSAWAVYV-----
KLH1-c : LSANINDHVTKEHSVPFNVFDYKTNFYD TLEFNGLSISLQNKLEAIKSDRFFAGFLLSGFKKSSLVKFNIC-----
Mth400-c : QTSTVNDPDSVTRDHSVPFDVFDYSSSFHYDYNLEFNGLSIAQLQHEVVRKAHERV FVAGFMLHGIKQALVRFEIC-----
Mth550-c : QTVSNPDPMTRDHSVPFEI FDYQSSPHYTYDLEFNGLSIPQLQHEVVRQAHERV FVAGFMLKGIKQALVTFEIC-----
KLH1-f : YE--INHQPTKKHAVPNDVFKY-ELLGYRYDNL EIGGMLHEI EKEIKDKQHHVRFVAGFLLHGI RPSADVQFQIC-----
Mth400-f : -SAKLD PDPTRIRSHANPSTVFN I-YDLGKYDYDQIGGKNLDEALIKDNQAHPRV FVAGFHLKGI GSSADVVFQICKD-----
Mth550-f : -SPDLDPDLEVRAHAVPNSVFN Y-FELGKYDNL EIGGSLEQLEDRI RQQSHARV FVAGFYLGGI GTSAGVTFAVCRK-----
    
```

```

*          380         *          400         *          420         *          440         *
KLH1-a : -----RNDHCEKAGDFV LGGPSEMKQFPRPYLFDLSDTVHKMGMKLDGCHYT-----VKAELFSVNGTALPDDL LPHVPPVVHH
Mth400-a : RTIGSADLDVEHHC EFAGDFLLGGANEM PWEHPLPFYFDVSDAVQRLGLSLTSKWS-----LEWDLYHVNGTALP GNL LPPYGKHR
Mth550-a : R-----DASDNCEFAGDFLLGGANEM PWEHPLPYFDVTEAVQRLGLVALDGAWT-----VKVDFYHVNGTAMP GDI LPPVGAHR
KLH1-b : -----TTDSVQHKACTFAVLGGSKEMKQFDRVFKFDITHLKLDLTDADGDFE-----VTVDIT EVDGTKLASSLIPHASVFI
Mth400-b : -----GSGGEPYLVKIAV LCGEKEMPHFDRLFKMEITDALSDGLSYDNDFD-----VTLQITDVTGKESVSDS FQSDTIFHE
Mth550-b : -----VAGGQRYRAGGVAVLGGEQEMPHFDRLFKMEITDALKALGLGYNDDF-----ISLNI TDVNGKHVDENTFSHDTIFHD
KLH1-c : -----TSSNCHPAC EGYLLGDENEM PWAYRVFVKYDITEKHLDKLHAE DHFY-----IDYEVFDLKPASLGKDLFKQPSVIHE
Mth400-c : -----SSGGACGQAGEFYILGDDYEM PWEYDRLFKYEITDQLKEQGLEPLDRYD-----IKYTVFALDGSVVG DTFGQVTVVHT
Mth550-c : -----QSGGQCSAAGEFYMLGDEYEM PWWYDRLFKYEITQQLADKGLKPLDYSY-----VQYQVKDLQNNVIGD NFGKPTIVHT
KLH1-f : -----KTS EDCDHGGQIFVLGGTKEMAWYDRLFKYDITHALHDAITPEDVPH-PSEPFPIKVS VTA VNGT VLPASILHAPT IYE
Mth400-f : -----KSGGQCTRAGSIFILGGEKEM PWEYDRLYKYDITDALKDNGLQPEDVFD-AEAPFY LKYEVTAVNGSTLPRSVINAPTLIFV
Mth550-f : -----DDGKRCY YGGSFIFLGGSKEM PWSYDRVYKHDMDRALHAI GLEPEDVFDYKKNVNFISYNVTVAVNNTALPRSIIPKPTILYE
    
```

- KLH1-a : PEKGFTD : 411
- Mth400-a : PAKGDRD : 423
- Mth550-a : PAKGAHD : 422
- KLH1-b : HAR : 415
- Mth400-b : PGK : 421
- Mth550-b : PPH : 415
- KLH1-c : PRI : 414
- Mth400-c : YGA : 412
- Mth550-c : YGA : 414
- KLH1-f : PGL : 417
- Mth400-f : PAT : 420
- Mth550-f : PAK : 421

|                                       |  |   |
|---------------------------------------|--|---|
| <span style="color: red;">■</span>    | sauer  | } Aminosäuren der direkten Kontaktstellen |
| <span style="color: cyan;">■</span>   | basisch  |   |
| <span style="color: yellow;">■</span> | unpolar/hydrophob  |   |
| <span style="color: green;">■</span>  | polar/neutral  |   |
| <span style="color: grey;">■</span>   | weitere an den Dekamer-Dekamer Interfaces beteiligte Aminosäuren |   |
| <span style="color: black;">■</span>  | hochkonservierte Reste (100%)                                    |   |

**Abb. 51: Sequenzalignment der FU-a, FU-b, FU-c und FU-f mit Darstellung der an den Dekamer-Dekamer-Interfaces beteiligten Aminosäuren**

Die FU-a, FU-b, FU-c und FU-f der Hämocyane von *Megathura crenulata* (KLH1) und *Melanoides tuberculata* (MtH<sub>400</sub> und MtH<sub>550</sub>) wurden mit Hilfe des Programms MAFFT unter Verwendung der BLOSUM 62 Matrix aligniert. Das resultierende Sequenzalignment wurde mit der Software GENEDOC illustriert. Die an den Dekamer-Dekamer-Interfaces des KLH1 beteiligten sowie die bei den Untereinheiten des MtH an den entsprechenden Positionen vorhandenen Aminosäuren (Gatsogiannis & Markl, 2009) sind entsprechend ihrer physikalisch-chemischen Eigenschaften farblich gekennzeichnet (siehe Abbildungslegende).

Ein weiteres Dekamer-Dekamer-Interface wurde zwischen den funktionellen Einheiten „c<sub>2</sub>“ und „a<sub>2</sub>“ beschrieben. Zum einen spielen hier ionische Wechselwirkungen zwischen einem Glutamat (28) und einem Arginin (98) eine Rolle und zum anderen scheinen die zwei positiv geladenen Aminosäuren Arginin (100) und Lysin (35) potentielle Bindungspartner für ein brückenbildendes divalentes Anion zu sein (Gatsogiannis & Markl, 2009). Bei beiden Untereinheiten des MtH sind die Ladungsverhältnisse an diesen Kontaktstellen durch Aminosäuresubstitutionen so verändert, dass ein Interface in der für KLH1 beschriebenen Art hier nicht plausibel erscheint.

Keines der an der Bildung der Didekamere beteiligten Interfaces (Gatsogiannis & Markl, 2009) konnte bei beiden Untereinheiten des MtH in der für KLH1 beschriebenen Art vorgefunden werden. Durch verschiedene Aminosäuresubstitutionen sind die Ladungsverhältnisse an den untersuchten Kontaktstellen so verändert, dass die im Fall des KLH1 beschriebenen Bindungsmechanismen bei MtH unmöglich erscheinen. Da die 400 kDa Untereinheit des MtH sowohl in der Lage ist Homo-Didekamere als auch Mega-Di- und Tridekamere zu bilden (Lieb *et al.*, 2010), muss auch das MtH über entsprechende Kontaktstellen verfügen.

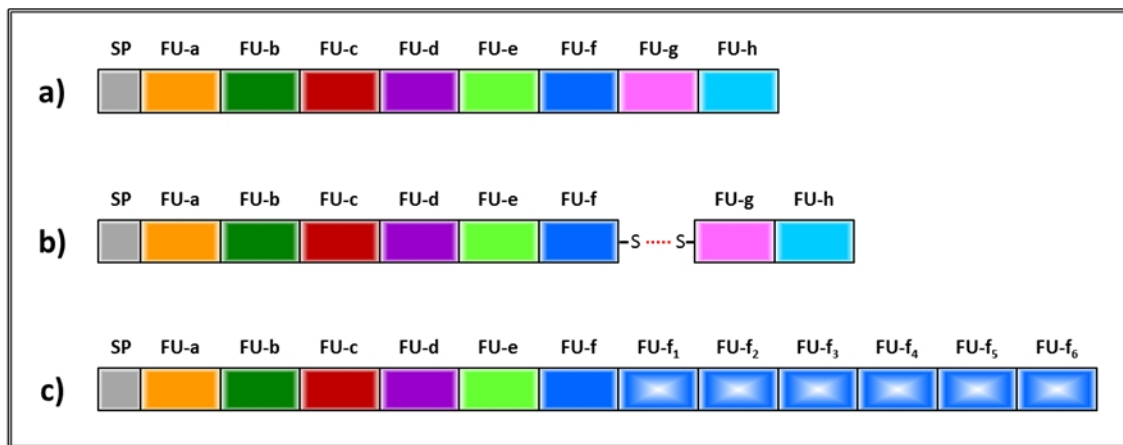
## 4.5 Evolution des Mega-Dekamers

Man geht davon aus, dass sich das Mollusken-Hämocyanin aus einem monomeren Vorläuferprotein entwickelt hat. Dieser Vorläufer muss bereits über die beiden Kupferbindungsstellen „Cu-A“ und „Cu-B“ verfügt haben (van Holde & Miller, 1995; Miller *et al.*, 1998; Lieb *et al.*, 2000; van Holde *et al.*, 2001). Es erfolgten offenbar drei Exonduplikations- und Fusionsereignisse, die zur Bildung der aus den acht funktionellen Einheiten FU-a bis FU-h bestehenden Hämocyanin-Untereinheit führten (Lieb *et al.*, 2001; Lieb & Markl, 2004). Das nur sieben FUs umfassende Hämocyanin der meisten Cephalopoden entwickelte sich dann durch den Verlust einer FU aus der aus acht FUs bestehenden Vorläufer-Untereinheit (Lieb *et al.*, 2000; Lieb *et al.*, 2001).

Das Mega-Hämocyanin, bei dem es sich höchstwahrscheinlich um eine Synapomorphie der Cerithioidea handelt (Lieb *et al.*, 2010), muss sich also auch aus einer acht FUs umfassenden frühen Untereinheit entwickelt haben. Das bedeutet, dass es im Fall der MtH<sub>550</sub>-Untereinheit zu Beginn zu einem Verlust der C-terminalen FU-g und FU-h gekommen sein muss. Erst danach konnten sich die sechs zusätzlichen FU-f der 550 kDa Untereinheit durch mehrere Exonduplikationen und Fusionen entwickeln.

Bei Untersuchungen am Hämocyanin der marinen Schnecke *Rapana thomasiana* (RtH2) konnte eine inter-FU-Disulfidbrücke zwischen den funktionellen Einheiten „f“ und „g“ identifiziert werden (Abb. 52b) (Gebauer *et al.*, 1999). Dieses höchst ungewöhnliche Strukturelement wurde bisher bei keinem anderen Mollusken-Hämocyanin gefunden.

Unter reduzierenden Bedingungen oder nach längerer Lagerung konnte die Spaltung der RtH2-Untereinheit in die beiden Bruchstücke RtH2.1 und RtH2.2 beobachtet werden. Weitere Untersuchungen zeigten, dass RtH2.1 die sechs an der Bildung der Zylinderwand beteiligten FU-a bis FU-f beinhaltet, während RtH2.2 aus den beiden kragenbildenden FU-g und FU-h bestand (Gebauer *et al.*, 1999). Somit verfügt das Hämocyanin von *Rapana thomasiana* (RtH2) über eine „Sollbruchstelle“ zwischen den funktionellen Einheiten „f“ und „g“.



**Abb. 52:** Schema der Hämocyanin-Untereinheiten von *Melanoides tuberculata* und *Rapana thomasiana* (RtH2)

Die Abbildung zeigt die schematische Darstellung der Untereinheiten des Hämocyanins von *Melanoides tuberculata* (MtH<sub>400</sub>, MtH<sub>550</sub>) sowie *Rapana thomasiana* (RtH2). Im Gegensatz zur 400 kDa Untereinheit des MtH (**a**) verfügt die RtH2-Untereinheit (**b**) über eine Disulfidbrücke zwischen den FU-f und FU-g. Diese „Sollbruchstelle“ könnte bei einer Vorläuferform der 550 kDa Untereinheit des MtH die Abspaltung der FU-g und FU-h erleichtert und die Duplikationen der FU-f erst ermöglicht haben (**c**).

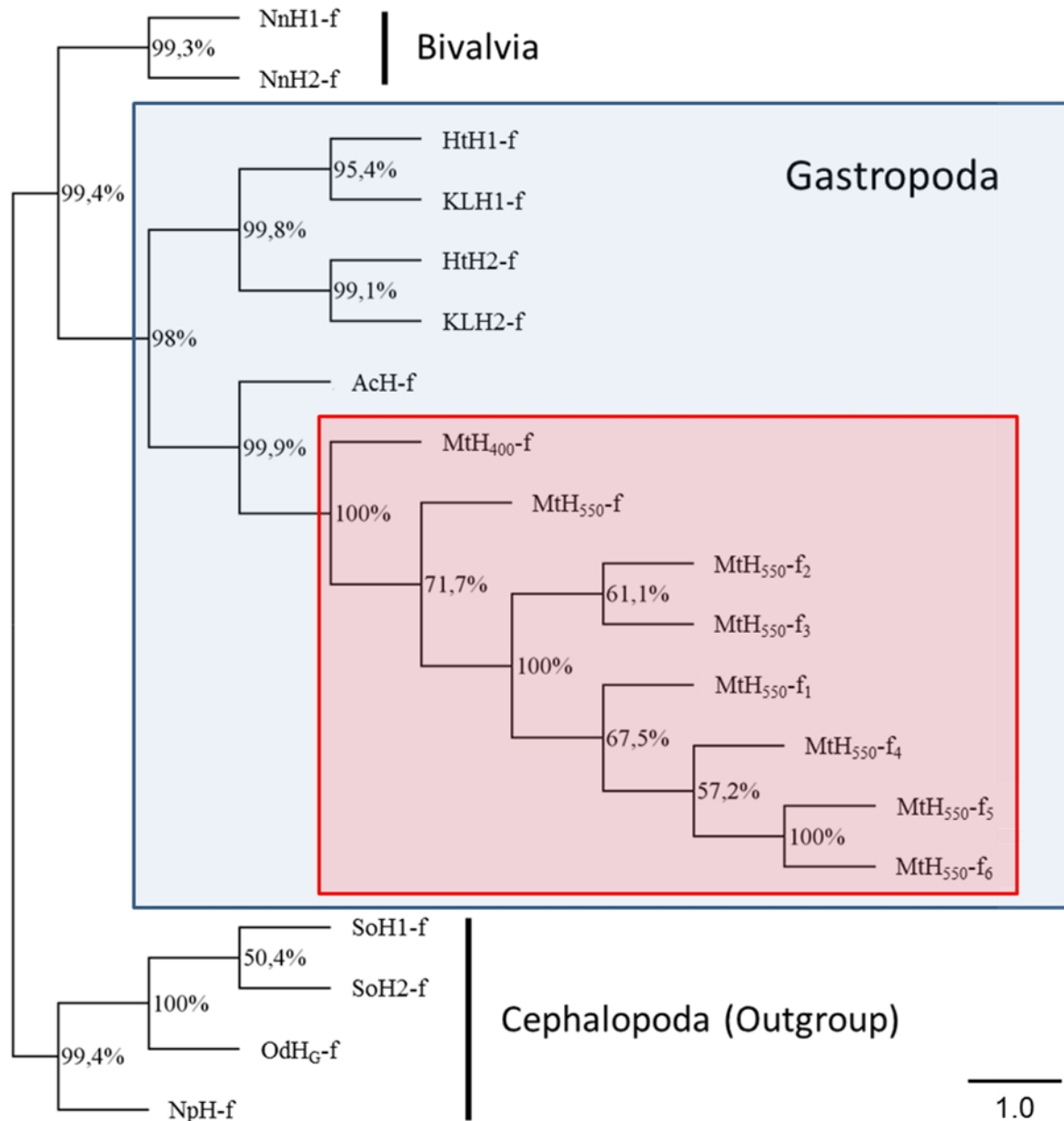
Die Gattung *Rapana* gehört zur Gruppe der Muricoidea innerhalb der Hypsogastropoda. Diese werden ebenso wie die Cerithioidea den Sorbeoconcha bzw. den Caenogastropoda

zugeordnet (siehe auch Abb. 8). Aufgrund der relativ kurzen phylogenetischen Distanz zwischen den beiden Familien ist es also denkbar, dass bereits ein gemeinsamer Vorläufer der Muricoidea und Cerithioidea die Disulfidbrücke zwischen den FU-f und FU-g aufgewiesen hat. Dieser Umstand hätte im Fall der Cerithioidea den Verlust der FU-g und FU-h erleichtert und die Entwicklung der zusätzlichen FU-f Varianten ermöglicht.

#### 4.5.1 Entstehung der sechs C-terminalen FU-f Varianten

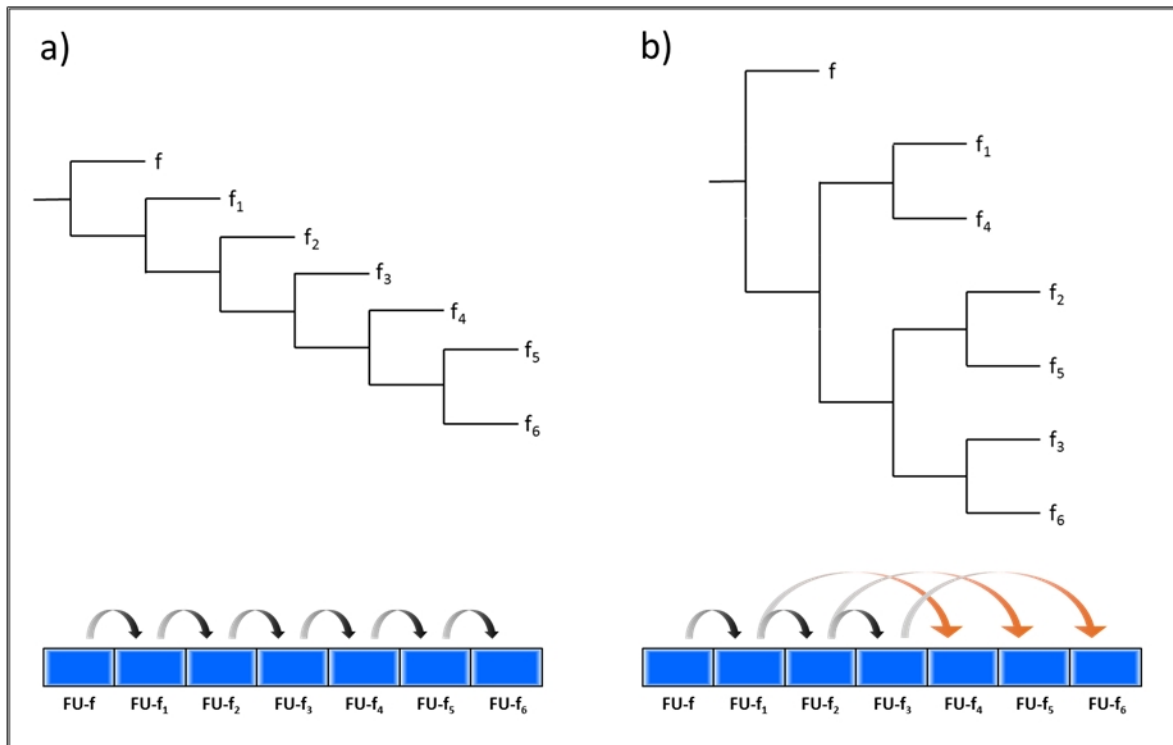
Gen- bzw. Exonduplikationen spielen bei der Evolution eukaryotischer Gene eine wichtige Rolle (Ohno, 1970; Letunic *et al.*, 2002; Zhang, 2003). Diese Duplikationsereignisse waren vermutlich auch für die Entwicklung der acht FUs umfassenden Untereinheit des Mollusken-Hämocyanins aus einer einzelnen Vorläufer-FU verantwortlich (Lieb *et al.*, 2001; Lieb & Markl; 2004). Die sechs zusätzlichen FUs der 550 kDa Untereinheit des MtH lassen sich eindeutig der FU-f zuordnen (Abb. 30). Sie müssen daher ebenfalls durch mehrere Exonduplikationen und Fusionen aus der „eigentlichen“, an der Bildung der Zylinderwand beteiligten FU-f entstanden sein.

Um die Entwicklung der sechs zusätzlichen FU-f Varianten der MtH<sub>550</sub>-Untereinheit aufzuklären, wurden sowohl die Nukleinsäure- als auch die abgeleiteten Aminosäuresequenzen der FU-f des MtH mit den FU-f anderer Mollusken-Hämocyanine aligniert. Die resultierenden Alignments wurden für die phylogenetische Rekonstruktion der Duplikationsereignisse der FU-f verwendet. Dafür wurden die Nukleinsäurealignments mit Hilfe der *Bayesianischen*-Analyse (MRBAYES) sowie der *Maximum-Likelihood* Analyse (RAXML, PHYML, MEGA5), jeweils auch unter Ausschluss der dritten Codonposition untersucht und die berechneten Stammbäume verglichen. Auch mit den Aminosäurealignments wurden sowohl *Bayesianische*- als auch *Maximum-Likelihood*-Analysen durchgeführt. Alle erhaltenen Stammbäume zeigten die in Abb. 53 dargestellte Topologie mit ähnlichen bis deutlich niedrigeren Unterstützungswerten. Aufgrund der teils sehr schlechten Unterstützungswerte kann unter Verwendung der zur Zeit vorhandenen Daten keine definitive Aussage über den Ablauf der Duplikationsereignisse getroffen werden. Es sind jedoch verschiedene Szenarien denkbar, die im Folgenden diskutiert werden. Eine serielle Duplikation der einzelnen FU-f (Abb. 54a) ist durch die erhaltene Topologie der phylogenetischen Rekonstruktion ebenso ausschließbar, wie eine Duplikation in Form von zwei Dreierblöcken (Abb. 54b). Die in diesen beiden Fällen zu erwartende Aufspaltung (Abb. 54) weicht deutlich von der durch die Berechnungen (Abb. 53) erhaltenen ab.



**Abb. 53: Phylogenetische Analyse der FU-f des Mth**

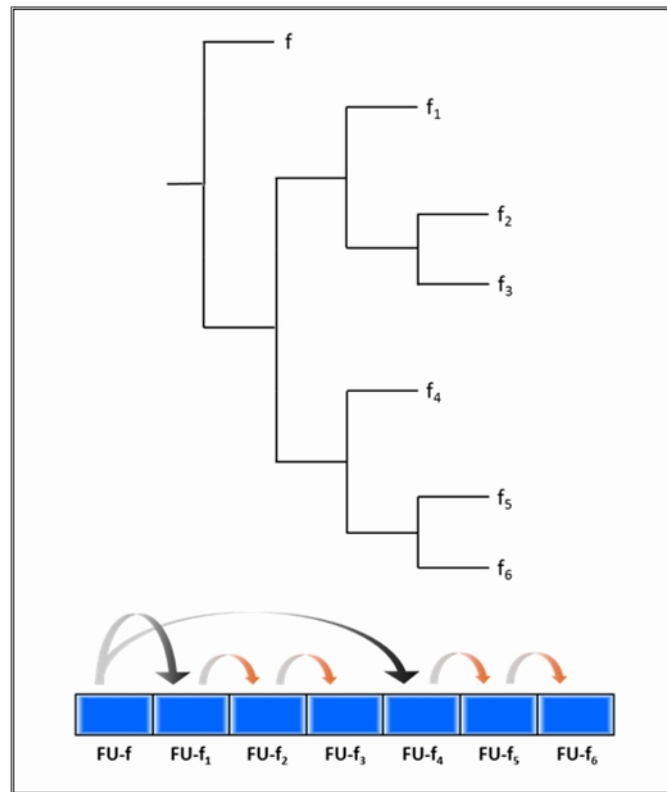
Die Abbildung zeigt den Konsensusbaum einer *Maximum-Likelihood* Analyse (RAXML; Stamatakis, 2006a), welcher auf einem 1317 Positionen umfassenden Nukleinsäurealignment (MAFFT; Katoh *et al.*, 2002) der FU-f von *Megathura crenulata* (KLH1, KLH2), *Haliotis tuberculata* (HtH1, HtH2), *Aplysia californica* (AcH), *Nucula nucleus* (NnH1, NnH2), *Nautilus pompilius* (NpH), *Octopus dofleini* (OdH), *Sepia officinalis* (SoH1, SoH2) sowie *Melanoides tuberculata* (MtH<sub>400</sub>, MtH<sub>550</sub>) basiert. Die Auswahl des geeigneten Evolutionsmodells (GTR+ $\Gamma_6$ +I) erfolgte mit der Software MEGA5 (Tamura *et al.*, 2011). Zur statistischen Absicherung wurden 10.000 *Bootstrap*-Replikate berechnet.



**Abb. 54: Hypothetische Topologien der Stammbäume bei zwei verschiedenen Duplikationsabläufen**

Bei einer seriellen Duplikation der FU-f, bei der die FUs einzeln nacheinander entstehen, wäre die in (a) gezeigte Topologie der phylogenetischen Rekonstruktion zu erwarten. Das entsprechende Ablaufschema der Duplikationen zeigt die untere Grafik in (a). Bei einer Duplikation in Dreierblöcken käme es zuerst zur Entwicklung der FU-f<sub>1</sub> bis FU-f<sub>3</sub> durch serielle Duplikationen. Danach würde die Duplikation des kompletten Blocks die sechs zusätzlichen FU-f<sub>1</sub> bis FU-f<sub>6</sub> ergeben (b, unten). Die dabei zu erwartende Topologie der phylogenetischen Rekonstruktion zeigt der obere Bereich von (b). Schwarze Pfeile kennzeichnen die frühen, orangefarbene Pfeile die späten Duplikationen.

Ein wahrscheinlicheres Szenario für die Entwicklung der FU-f<sub>1</sub> bis FU-f<sub>6</sub> zeigt Abb. 55. Hier kam es initial zur Bildung der FU-f<sub>1</sub> und FU-f<sub>4</sub>. Anschließend führten weitere Genduplikationen zur Entwicklung der FU-f<sub>2</sub> und FU-f<sub>3</sub> bzw. der FU-f<sub>5</sub> und FU-f<sub>6</sub>. Die hypothetische Stammbaumtopologie dieses Szenarios zeigt deutliche Gemeinsamkeiten mit dem tatsächlich berechneten Stammbaum in Abb. 53. Die FU-f<sub>2</sub> und FU-f<sub>3</sub> treten hier ebenso wie die FU-f<sub>5</sub> und FU-f<sub>6</sub> als monophyletische Gruppen auf. Im Fall der beiden letztgenannten FUs wurde dies in allen durchgeführten Berechnungen mit einem Unterstützungswert von 100 % bestätigt. Auch die im Bezug zu den FU-f<sub>5</sub> und FU-f<sub>6</sub> basale Stellung der FU-f<sub>4</sub> wurde in allen errechneten Stammbäumen gezeigt. Einzig die Position der FU-f<sub>1</sub>, welche bei den berechneten Stammbäumen basal zur FU-f<sub>4</sub> steht (Abb. 53) weicht von der hypothetischen Topologie (Abb. 55) ab.

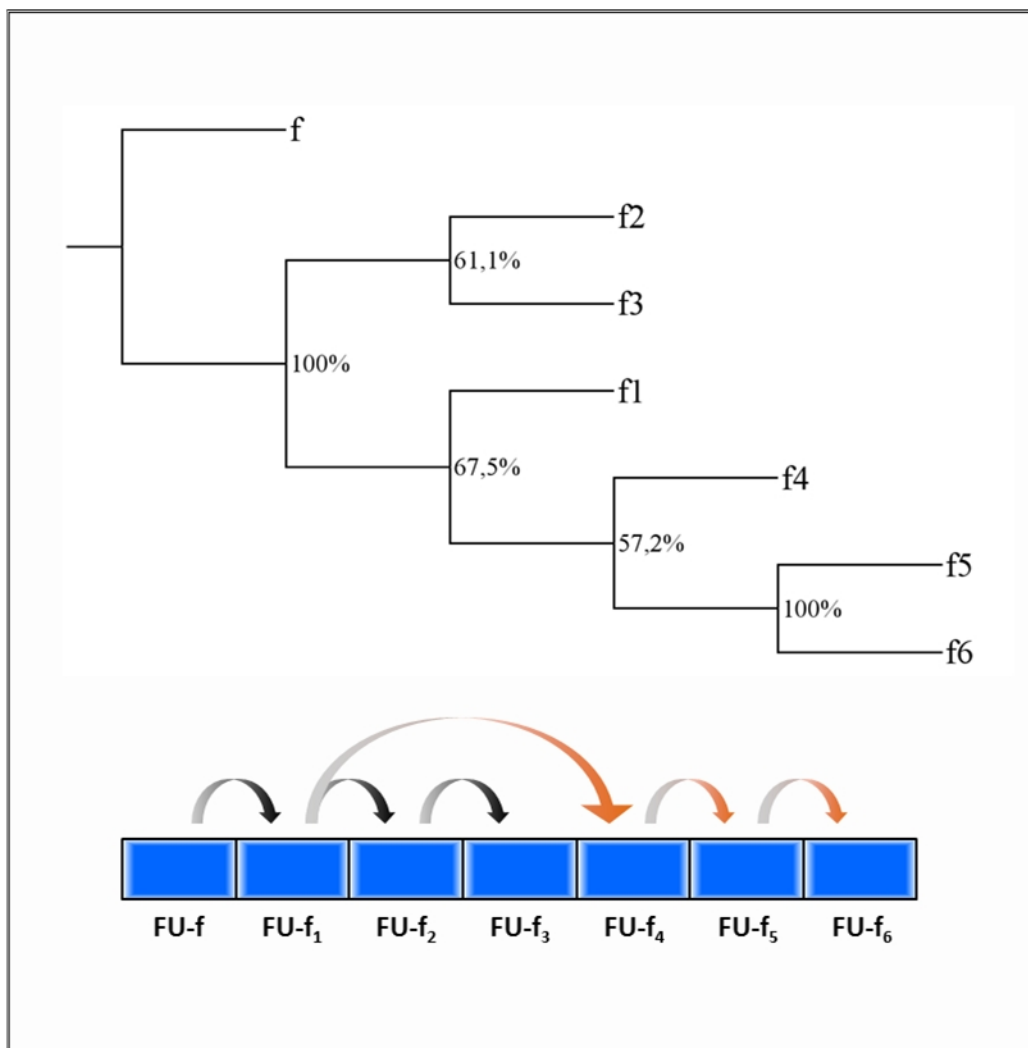


**Abb. 55: Hypothetische Stammbaumtopologie bei früher Entwicklung der FU-f<sub>1</sub> und FU-f<sub>4</sub>**

Die Abbildung zeigt einen hypothetischen Duplikationsablauf, bei dem zuerst die FU-f<sub>1</sub> und FU-f<sub>4</sub> entstehen. Danach entwickeln sich aus ihnen die FU-f<sub>2</sub> und FU-f<sub>3</sub> bzw. die FU-f<sub>5</sub> und FU-f<sub>6</sub> (unten). Schwarze Pfeile zeigen die frühen, orangefarbene Pfeile die späten Duplikationsereignisse. Der obere Teil der Abbildung zeigt die dazugehörige Topologie der phylogenetischen Rekonstruktion.

Eine weitere Hypothese für die Entwicklung der sechs FU-f, welche auf der tatsächlich gefundenen Topologie der berechneten Stammbäume beruht, zeigt Abb. 56. Deutlich erkennbar sind auch hier die beiden monophyletischen Gruppen der FU-f<sub>2</sub> und FU-f<sub>3</sub> bzw. der FU-f<sub>5</sub> und FU-f<sub>6</sub>. Mehrere Genduplikationsereignisse führten initial zur Bildung der FU-f<sub>1</sub>, FU-f<sub>2</sub> und FU-f<sub>3</sub>. Durch eine weitere Duplikation der FU-f<sub>1</sub> bildete sich schließlich die FU-f<sub>4</sub>. Dabei muss es sich um eine Transduktion der „originalen“ FU-f<sub>1</sub> gehandelt haben, wie sie z.B. durch Retrotransposons des „Non-LTR-Typs“ (Retroposons) möglich ist. Im Falle des humanen Retroposons „L1“ wurde eine Transduktion des flankierenden 3'-Sequenzabschnittes an eine andere genomische Position mehrfach beschrieben (Moran *et al.*, 1999; Goodier *et al.*, 2000; Ostertag *et al.*, 2001). Auch für weitere eukaryotische Transposons konnte die unbalancierte Translokation eines nicht repetitiven Sequenzabschnittes an eine andere Position im Genom nachgewiesen werden (Zelnick *et al.*, 1987; Bailey *et al.*, 2003). Man geht heute davon aus, dass mobile DNA-Elemente eine wichtige Rolle bei der Evolution des eukaryotischen Genoms gespielt haben und weiterhin spielen. Gerade für die Bildung neuer Gene sowie für das sogenannte „domain-shuffling“

scheinen transposable Elemente essentiell zu sein (Bowen *et al.*, 2002; Kazazian, 2004; Babushok *et al.*, 2007; Feschotte *et al.*, 2007).



**Abb. 56: Topologie der FU-f der 500 kDa Untereinheit des Mth**

Die Abbildung zeigt die Topologie der FU-f der Mth<sub>550</sub>-Untereinheit basierend auf den berechneten Stammbäumen. Die Unterstützungswerte resultieren aus 10.000 *Bootstrap*-Replikaten, die im Rahmen einer *Maximum-Likelihood* Analyse (RAXML) des Nukleinsäurealignments (MAFFT) der FU-f berechnet wurden. Die phylogenetische Rekonstruktion wurde unter Verwendung des GTR+ $\Gamma_6$ +I Modells durchgeführt. Für das Alignment wurde eine PAM100-Matrix verwendet. Der untere Teil der Abbildung zeigt den hypothetischen Ablauf der Duplikationsereignisse, die zu der gefundenen Topologie geführt haben könnten. Schwarze Pfeile kennzeichnen die frühen, orangefarbene Pfeile die späten Duplikationen.

Im Falle der Duplikation der FU-f<sub>4</sub> des Mega-Hämocyanins von *Melanoides tuberculata* muss es sich um eine Transposition nach dem „Copy-Paste“-Mechanismus, bei dem eine Kopie des ursprünglichen Sequenzabschnittes an einer anderen Stelle im Genom eingefügt wird, gehandelt haben. Dies wird auch durch die hohe Sequenzidentität von 66 % zwischen den FU-f<sub>1</sub> und FU-f<sub>4</sub> untermauert (siehe Tab.1). Weitere Duplikationen der FU-f<sub>4</sub> führten anschließend zur Bildung der C-terminalen FU-f<sub>5</sub> und FU-f<sub>6</sub>.

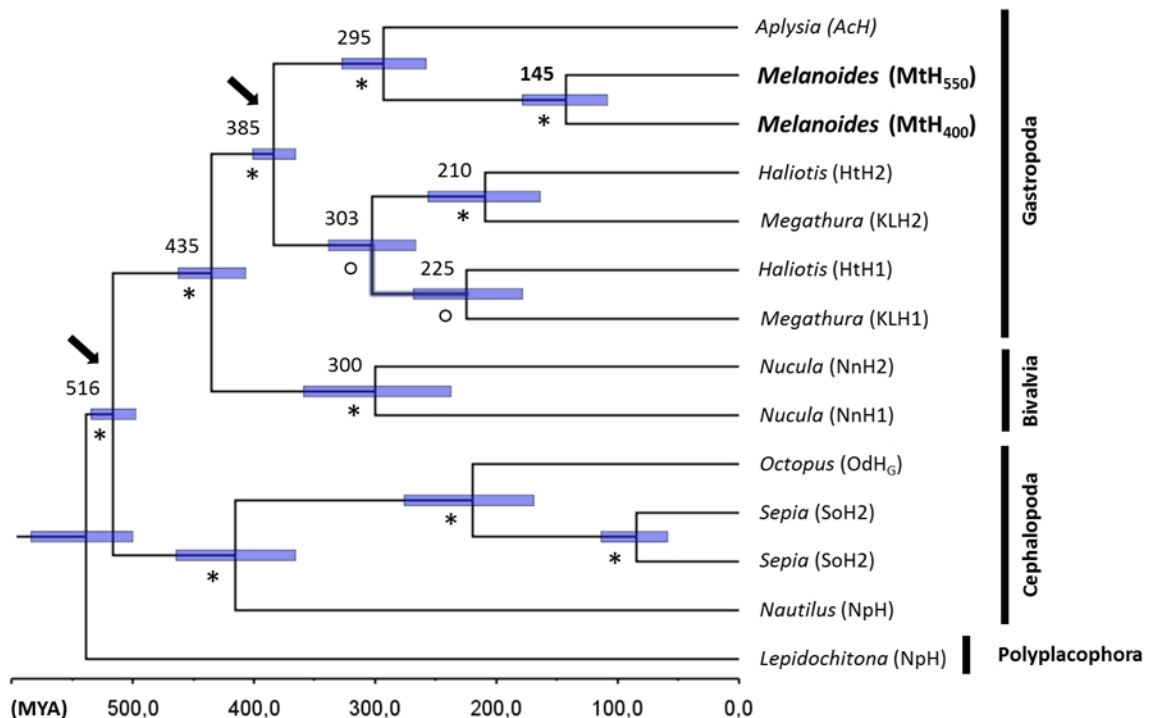
## 4.6 Molekulare Uhr

Um die Entstehung des Mega-Hämocyanins der Cerithioidea datieren zu können, wurden die konkatenierten Nukleotidsequenzen der FU-a bis FU-f der Hämocyanine von *Nucula nucleus* (NnH1, NnH2), *Aplysia californica* (AcH), *Haliotis tuberculata* (HtH1, HtH2), *Megathura crenulata* (KLH1, KLH2), *Nautilus pompilius* (NpH), *Octopus dofleini* (OdHG), *Sepia officinalis* (SoH1, SoH2) sowie *Melanoides tuberculata* (MtH<sub>400</sub>, MtH<sub>550</sub>) verwendet, um eine molekulare Uhr zu berechnen. Mit Hilfe des Programms MAFFT (Katoh *et al.*, 2002) wurde ein Sequenzalignment erstellt und anschließend manuell editiert und bereinigt. Die nur bei SoH1 und SoH2 vorhandene FU-d<sup>c</sup> wurde bei den Berechnungen nicht berücksichtigt und aus dem Alignment entfernt. Für die Auswahl eines geeigneten Substitutionsmodells wurde ebenso wie für die anschließende Berechnung eines initialen *Maximum Likelihood* Stammbaumes (GTR+ $\Gamma_4$ +I; 10000 *Bootstrap*-Replikate) das Programm MEGA5 verwendet (Tamura *et al.*, 2011). Die Berechnung der molekularen Uhr erfolgte durch eine *bayesianische* MCMC-Analyse mit dem Programm BEAST 1.6.2 (Drummond *et al.*, 2007) unter Verwendung von zwei getrennten Codon-Partitionen [(1+2), 3]. Für die Datierung des Mega-Hämocyanins wurde eine „*relaxed clock*“ mit einer nicht korrigierten *lognormalen* Verteilung gewählt. Als „*Tree-Prior*“ wurde das „*Yule-Modell*“ verwendet und insgesamt 50 Millionen Generationen berechnet. Dabei wurden die Kettenzustände alle 1000 Generationen gespeichert. Nach Abschluss der Berechnungen wurde die Konvergenz aller Parameter des Datensatzes unter Verwendung des Programms TRACER (Rambaut *et al.*, 2007) überprüft. Die ersten 12500 berechneten Bäume wurden ignoriert (*burnin*: 25 %).

Für die Kalibrierung der berechneten molekularen Uhr wurden zwei verschiedene, durch Fossilienfunde belegte Zeitpunkte gewählt. Cephalopoden und Gastropoden traten erstmalig gemeinsam im späten Kambrium (Merioneth bzw. Furongian; 501 bis 488 MYA) auf (Benton, 1993; Gradstein *et al.*, 2004). Dieser Zeitabschnitt wurde als Eichpunkt für die Trennung dieser beiden Molluskenklassen verwendet. Als zweiter Eichpunkt diente die Trennung der Heterobranchia von den Vetigastropoda, die vor etwa 380 Millionen Jahren stattfand (Frýda *et al.*, 2001).

Das resultierende Chronogramm (Abb. 57) zeigt eine an allen Verzweigungen gut unterstützte Topologie (pp  $\geq$  0,98 bzw. bs  $\geq$  98 %). Innerhalb der Cephalopoden stimmen die errechneten Zeitpunkte sehr gut mit aktuell publizierten, auf Hämocyanin-Daten basierenden molekularen Uhren überein (Warnke *et al.*, 2011). Sowohl die Entwicklung der beiden Isoformen des SoH ( $85 \pm 25$  MYA) als auch die Trennung der Decabrachia und

Octabrachia ( $220 \pm 50$  MYA) wurden bereits mehrfach entsprechend datiert (Lieb & Markl, 2004; Bergmann *et al.*, 2006; Warnke *et al.*, 2011). Die Trennung der Dibranchiata und Tetrabranchiata datiert die berechnete molekulare Uhr auf  $415 \pm 45$  MYA zurück. Dieses Alter stimmt mit den Berechnungen von Lieb & Markl (2004) und Bergmann *et al.* (2006, 2007) überein. Warnke *et al.* (2011) datieren dieses Ereignis auf  $450 \pm 60$  MYA zurück. Sowohl die Nautilida als auch die Bactritoidea, die als gemeinsamer Vorläufer der Ammoniten und Coleoiden gelten (Chen *et al.*, 1983; Nishiguchi *et al.*, 2008), traten erstmals an der Grenze des Silurs zum Devon auf (ca. 416 MYA) (Benton, 1993; Nishiguchi *et al.*, 2008). Diese Daten stimmen deutlich besser mit dem hier berechneten Trennungszeitpunkt ( $415 \pm 45$  MYA) überein.



**Abb. 57: Evolution der Mollusken-Hämocyanine (Molekulare Uhr)**

Die dargestellte molekulare Uhr basiert auf einem multiplen Sequenzalignment (MAFFT) der Wand-FUs (FU-a bis FU-f) der gezeigten Mollusken-Hämocyanine und wurde mittels des BEAST 1.6.2 Programmpaketes berechnet. Die Balken (blau) zeigen das „95 % posterior density interval“ für den entsprechenden Knoten. Als Kalibrierungspunkte (Pfeile) wurden die Trennung der Vetigastropoda und Heterobranchia ( $380 \pm 10$  MYA, Frýda *et al.*, 2001) sowie die Trennung der Cephalopoda und Gastropoda ( $500 \pm 10$  MYA, Benton, 1993) verwendet. Die Entstehung der beiden Untereinheiten des *Melanoides tuberculata* Hämocyanins fand vor etwa  $145 \pm 35$  MYA statt. Die Topologie des gezeigten Chronogramms wurde zusätzlich durch *Maximum-Likelihood* Analysen (RAXML) bestätigt und abgesichert (nicht gezeigt). Sterne kennzeichnen Verzweigungen mit *Bootstrap*- bzw. *posterior probability* Werten von 100 % bzw. 1.0. Kreise kennzeichnen Verzweigungen mit *Bootstrap*- bzw. *posterior probability* Werten von  $\geq 98$  % bzw.  $\geq 0,98$ .

Die Trennung der Protobranchia und Gastropoda datiert die molekulare Uhr auf  $435 \pm 35$  MYA zurück. Diese Daten entsprechen den durch Warnke *et al.* (2011) publizierten. Sowohl Lieb & Markl (2004) als auch Bergmann *et al.* (2007) setzen den entsprechenden Trennungszeitpunkt deutlich früher (480 bzw. 494 MYA) an. Erste Fossilienfunde der Nuculidae stammen aus dem Givetian (385 – 392 MYA), welches einen Abschnitt des mittleren Devons darstellt (Benton, 1993). Die berechneten Daten stehen dementsprechend im Einklang mit den Fossilienfunden. Die Entstehung der beiden Isoformen des Hämocyanins von *Nucula nucleus* geschah in Übereinstimmung mit den Daten von Warnke *et al.* (2011) vor ca. 300 MYA. Auch diese Daten, die zudem ein hohes Konfidenzintervall aufweisen ( $\pm 60$  MYA), weichen deutlich von den durch Lieb & Markl (2004) und Bergmann *et al.* (2007) bestimmten Zeitpunkten (370 bzw. 396 MYA) ab.

Die Entstehung der beiden Isoformen der Hämocyanine von *Haliotis tuberculata* und *Megathura crenulata* wurde mit  $303 \pm 35$  MYA berechnet. Diese Daten stimmen deutlich sowohl mit den aktuellen Daten von Warnke *et al.* (2011) als auch mit den älteren Datierungen (Lieb *et al.*, 2000; Lieb *et al.*, 2004; Lieb & Markl 2004; Bergmann *et al.*, 2006; 2007) überein. Für die Trennung der Arten *Haliotis* und *Megathura* wurde ein Alter von 210 bis 225 MYA berechnet.

Für die gezeigte molekulare Uhr (Abb. 57) wurden erstmals Daten von Caenogastropoden (MtH<sub>400</sub>, MtH<sub>550</sub>) in die Berechnungen einbezogen. Der vergrößerte Datensatz führt zu einer deutlich besseren Auflösung der Apogastropoden. Dies macht sich auch durch das relativ niedrige Konfidenzintervall ( $\pm 30$  MYA) bemerkbar. Die Trennung der Heterobranchia (AcH) von den Caenogastropoden (MtH) geschah laut den Berechnungen vor ca. 295 MYA. Weitere 150 MYA später entwickelten sich dann die beiden Untereinheiten des Hämocyanins von *Melanooides tuberculata*. Demnach handelt es sich bei der Evolution des Mega-Hämocyanins um ein geschichtlich relativ junges Ereignis. Der Ursprung der Überfamilie Cerithioidea wird in der Literatur kontrovers diskutiert und reicht vom Perm (299 - 251 MYA) bis zum Jura (201 – 145,5 MYA). In dem von Houbrick (1988) verfassten Artikel „Cerithioidean Phylogeny“ wurden die verschiedenen, deutlich voneinander abweichenden Datierungen ausgiebig diskutiert. Houbrick (1988) kommt zu dem Schluss, dass die angegebenen Daten für den Ursprung der meisten Familien der Cerithioidea deutlich zu früh angesetzt sind und dass die rezenten Familien anhand von Fossilienfunden nur bis in die späte Kreidezeit (99,6 – 65,5 MYA) nachvollzogen werden können. Diese Datierung steht in klarer Übereinstimmung mit den

Ergebnissen der berechneten molekularen Uhr (Abb. 57), die die Entstehung des Mega-Hämocyanins der rezenten Cerithioidea auf  $145 \pm 35$  MYA datiert.

Insgesamt gesehen stimmen die berechneten Zeitpunkte in großen Teilen mit den früheren, auf Berechnungen aus Distanzmatrizen basierenden Daten von Lieb & Markl (2004) und Bergmann *et al.* (2006) überein. Zudem entsprechen die hier gezeigten Daten sehr genau den aktuell publizierten Berechnungen von Warnke *et al.* (2011), denen ebenfalls eine *bayesianische* MCMC-Analyse unter Annahme einer „*relaxed-clock*“ zugrunde lag.

Das ständig zunehmende Fossilien-Repertoire und die wachsende Anzahl an verfügbaren Sequenzdaten werden in der Zukunft weitaus genauere und verlässlichere Berechnungen der Mollusken- und Hämocyanin-Evolution möglich machen. Zudem ist auch die Datierung der Erdzeitalter ständigen Veränderungen unterworfen. Alleine der Zeitpunkt des Kambrium-Ordovizium-Überganges schwankte in den seit 1982 veröffentlichten Zeitskalen um 22 MYA (Gradstein *et al.*, 2004). Auch diese Abweichungen führen natürlich zu Unterschieden zwischen den früheren und aktuellen Datierungen, werden sich aber auch in Zukunft nicht vollständig vermeiden lassen.

## 5 Zusammenfassung

Bei dem 2010 von unserer Arbeitsgruppe entdeckten Mega-Hämocyanin handelt es sich um einen stark abgewandelten Typ des respiratorischen Proteins Hämocyanin, bestehend aus zwei flankierenden regulären Dekameren und einem zentralen Mega-Dekamer. Diese sind aus zwei immunologisch verschiedenen Untereinheiten mit ~400 bzw. ~550 kDa aufgebaut, die in unserer Arbeitsgruppe bereits proteinbiochemisch charakterisiert wurden. Im Zuge dieser Untersuchungen konnte zudem eine 3D-Rekonstruktion des Oligomers (13,5 MDa) mit einer Auflösung von 13Å erstellt werden. Das Ziel der vorliegenden Arbeit war die Aufklärung der Primärstruktur beider Polypeptide bei der Schnecke *Melanoides tuberculata* (MtH). Es gelang, die cDNAs der beiden Untereinheiten vollständig zu sequenzieren. Die zu typischen Dekameren assemblierende MtH<sub>400</sub>-Untereinheit umfasst 3445 Aminosäuren und besitzt eine theoretische Molekularmasse von 390 kDa. Nach dem Signalpeptid von 23 Aminosäuren Länge folgen die für Gastropoden-Hämocyanine typischen funktionellen Einheiten FU-a bis FU-h. Insgesamt verfügt die MtH<sub>400</sub>-Untereinheit über sechs potentielle N-Glykosylierungsstellen. Die MtH<sub>550</sub>-Untereinheit, welche mit 10 Kopien das Mega-Dekamer bildet, umfasst 4999 Aminosäuren und besitzt eine theoretische Molekularmasse von 567 kDa. Damit handelt es sich bei dieser Untereinheit um die zweitgrößte jemals bei einem Protein detektierte Polypeptidkette. Die MtH<sub>550</sub>-Untereinheit besteht aus einem Signalpeptid von 20 Aminosäuren Länge und den typischen Wand-FUs (FU-a bis FU-f). Daran anschließend folgen sechs weitere Varianten der FU-f (FU-f<sub>1</sub> bis FU-f<sub>6</sub>). Die MtH<sub>550</sub>-Untereinheit verfügt über insgesamt zwölf potentielle N-Glykosylierungsstellen. Anhand der ermittelten Primärstrukturdaten wird klar, dass der auffällig vergrößerte Kragenbereich des Mega-Dekamers aus je 10 Kopien der FU-f<sub>1</sub> bis FU-f<sub>6</sub> besteht. Die ermittelten Sequenzdaten der beiden MtH-Untereinheiten weisen im Vergleich zu anderen Hämocyanin Sequenzen einige sehr charakteristische Indels sowie unübliche N-Glykosylierungsstellen auf. Es war zudem möglich, anhand einer molekularen Uhr den Entstehungszeitpunkt des Mega-Hämocyanins zu datieren ( $145 \pm 35$  MYA). Sowohl die Topologie als auch die berechneten Trennungszeitpunkte des an allen Verzweigungen gut unterstützten Stammbaums stimmen mit den bisher publizierten und auf Hämocyanindaten basierenden molekularen Uhren überein.

# Anhang

## A Abkürzungsverzeichnis

### A.1 Allgemeine Abkürzungen

|                    |   |
|--------------------|---|
| (m)M               | (milli)Molar  |
| (TEM)              | (Transmissions-) Elektronenmikroskopie                    |
| Å                  | Ångström  |
| Abb.               | Abbildung   |
| Amp                | Ampicillin  |
| AS                 | Aminosäure  |
| bp                 | Basenpaar   |
| cDNA               | „Copy“ oder „complementary“-DNA                           |
| Cu-A, Cu-B         | Kupferbindungszentren der Hämocyanine                     |
| Da                 | Dalton  |
| ddH <sub>2</sub> O | Bidestilliertes Wasser (H <sub>2</sub> O bidest)          |
| DEPC               | Diethylpyrocarbonat                                       |
| dH <sub>2</sub> O  | Destilliertes Wasser                                      |
| DMSO               | Dimethylsulfoxid  |
| DNA                | Desoxyribonukleinsäure                                    |
| dNTP               | Desoxynukleosidtriphosphat                                |
| DTT                | Dithiothreitol  |
| EDTA               | Ethylendiamintetraacetat                                  |
| EtBr               | Ethidiumbromid  |
| for                | <i>forward</i>  |
| FU                 | Functional Unit (funktionelle Einheit)                    |
| g                  | Erdbeschleunigung (Mittelwert: $g = 9,81 \text{ m/s}^2$ ) |

|                  |   |
|------------------|---|
| H <sub>2</sub> O | Wasser  |
| Hc               | Hämocyanin  |
| IPTG             | Isopropyl-β-D-1-Thiogalactopyranosid                    |
| JTT              | James-Taylor-Thornton                                   |
| Kan              | Kanamycin   |
| kb               | Kilobasen (1000 Nukleotide)                             |
| kDa              | Kilodalton (1000 Dalton)                                |
| LB               | lysogeny broth (Nährmedium)                             |
| MCS              | Multiple cloning site                                   |
| MOPS             | 3-(N-Morpholino)-2-Hydroxypropansulfonsäure Natriumsalz |
| mRNA             | Messenger RNA   |
| MW               | Molekulargewicht  |
| MYA              | Million years ago                                       |
| NaCl             | Natriumchlorid  |
| OD               | Optische Dichte   |
| ORF              | Open Reading Frame (offener Leserahmen)                 |
| PCR              | Polymerase chain reaction (Polymerasekettenreaktion)    |
| pI               | Isoelektrischer Punkt                                   |
| RACE             | Rapid amplification of cDNA ends                        |
| rev              | <i>reverse</i>  |
| RNA              | Ribonukleinsäure  |
| rpm              | Rounds per minute (Umdrehungen pro Minute)              |
| RT               | Reverse Transkription oder Raumtemperatur               |
| SP               | Signalpeptid  |
| TAE              | Tris-Acetat-EDTA  |
| Taq              | <i>Thermophilus aquaticus</i>                           |
| TBE              | Tris-Borat-EDTA   |
| TE               | Tris-EDTA   |

|       |   |
|-------|---|
| Tris  | Trishydroxymethylaminomethan                    |
| tRNA  | Transfer-RNA                                    |
| U     | Unit  |
| UE    | Untereinheit                                    |
| ÜN    | Über Nacht                                      |
| UTR   | Untranslated region (Untranslatierter Bereich)  |
| UV    | Ultraviolett                                    |
| v/v   | Volumen pro Volumen                             |
| w/v   | Gewicht pro Volumen                             |
| X-Gal | 5-Brom-4-Chlor-3-Indolyl- $\beta$ -D-Galactosid |

## A.2 Abkürzungen in der Hämocyaninforschung

| Hämocyanin-Abkürzung | Tierart                       | Klasse         |
|----------------------|-------------------------------|----------------|
| AcH                  | <i>Aplysia californica</i>    | Gastropoda     |
| HtH                  | <i>Haliotis tuberculata</i>   | Gastropoda     |
| KLH                  | <i>Megathura crenulata</i>    | Gastropoda     |
| LaH                  | <i>Leptochiton assellus</i>   | Polyplacophora |
| LcH                  | <i>Lepidochitona cinerea</i>  | Polyplacophora |
| LhH                  | <i>Laevipilina hyalina</i>    | Monoplacophora |
| MtH                  | <i>Melanoides tuberculata</i> | Gastropoda     |
| NnH                  | <i>Nucula nucleus</i>         | Bivalvia       |
| NpH                  | <i>Nautilus pompilius</i>     | Cephalopoda    |
| OdH                  | <i>Octopus dofleini</i>       | Cephalopoda    |
| RtH                  | <i>Rapana thomasi</i>         | Gastropoda     |
| SoH                  | <i>Sepia officinalis</i>      | Cephalopoda    |

## B IUPAC-Nomenklatur für Aminosäuren

| Trivialname | Abkürzung | Symbol |
|-------------|-----------|--------|
| Alanin      | Ala       | A      |
| Arginin     | Arg       | R      |
| Asparagin   | Asn       | N      |
| Aspartat    | Asp       | D      |
| Cystein     | Cys       | C      |
| Glutamat    | Glu       | E      |
| Glutamin    | Gln       | Q      |
| Glycin      | Gly       | G      |
| Histidin    | His       | H      |
| Isoleucin   | Ile       | I      |

| Trivialname  | Abkürzung | Symbol |
|--------------|-----------|--------|
| Leucin       | Leu       | L      |
| Lysin        | Lys       | K      |
| Methionin    | Met       | M      |
| Prolin       | Pro       | P      |
| Phenylalanin | Phe       | F      |
| Serin        | Ser       | S      |
| Threonin     | Thr       | T      |
| Tryptophan   | Trp       | W      |
| Tyrosin      | Tyr       | Y      |
| Valin        | Val       | V      |

## C IUPAC-Nomenklatur für Nukleotide

| Symbol | Nukleotid           |
|--------|---------------------|
| A      | Adenin              |
| B      | nicht Adenin        |
| C      | Cytosin             |
| D      | nicht Cytosin       |
| G      | Guanin              |
| H      | nicht Guanin        |
| K      | Guanin oder Thymin  |
| M      | Adenin oder Cytosin |

| Symbol | Nukleotid               |
|--------|-------------------------|
| N      | alle Nukleotide möglich |
| R      | Adenin oder Guanin      |
| S      | Cytosin oder Guanin     |
| T      | Thymin                  |
| U      | Uracil                  |
| V      | nicht Thymin            |
| W      | Adenin oder Thymin      |
| Y      | Cytosin oder Thymin     |

## D Der genetische Code

Die Tabelle zeigt alle 64 möglichen Codons und die dadurch kodierten Aminosäuren. Die unterschiedlichen Farben kennzeichnen die physikalisch-chemischen Eigenschaften der verschiedenen Aminosäuren: orange (unpolar / hydrophob), grün (polar / neutral), rot (sauer) sowie blau (basisch).

|         |   | 2. Base |              |     |          |     |           |     |            |
|---------|---|---------|--------------|-----|----------|-----|-----------|-----|------------|
|         |   | U       |              | C   |          | A   |           | G   |            |
| 1. Base | U | UUU     | Phenylalanin | UCU | Serin    | UAU | Tyrosin   | UGU | Cystein    |
|         |   | UUC     | Phenylalanin | UCC | Serin    | UAC | Tyrosin   | UGC | Cystein    |
|         |   | UUA     | Leucin       | UCA | Serin    | UAA | STOP      | UGA | STOP       |
|         |   | UUG     | Leucin       | UCG | Serin    | UAG | STOP      | UGG | Tryptophan |
|         | C | CUU     | Leucin       | CCU | Prolin   | CAU | Histidin  | CGU | Arginin    |
|         |   | CUC     | Leucin       | CCC | Prolin   | CAC | Histidin  | CGC | Arginin    |
|         |   | CUA     | Leucin       | CCA | Prolin   | CAA | Glutamin  | CGA | Arginin    |
|         |   | CUG     | Leucin       | CCG | Prolin   | CAG | Glutamin  | CGG | Arginin    |
|         | A | AUU     | Isoleucin    | ACU | Threonin | AAU | Asparagin | AGU | Serin      |
|         |   | AUC     | Isoleucin    | ACC | Threonin | AAC | Asparagin | AGC | Serin      |
|         |   | AUA     | Isoleucin    | ACA | Threonin | AAA | Lysin     | AGA | Arginin    |
|         |   | AUG     | Methionin    | ACG | Threonin | AAG | Lysin     | AGG | Arginin    |
|         | G | GUU     | Valin        | GCU | Alanin   | GAU | Aspartat  | GGU | Glycin     |
|         |   | GUC     | Valin        | GCC | Alanin   | GAC | Aspartat  | GGC | Glycin     |
|         |   | GUA     | Valin        | GCA | Alanin   | GAA | Glutamat  | GGA | Glycin     |
|         |   | GUG     | Valin        | GCG | Alanin   | GAG | Glutamat  | GGG | Glycin     |

## E Programme

| Programm                | Internet-Adresse  |
|-------------------------|---|
| BEAST 1.6.2             | <a href="http://beast.bio.ed.ac.uk/Main_Page">http://beast.bio.ed.ac.uk/Main_Page</a>                                   |
| CLUSTALX / CLUSTALW     | <a href="http://www.clustal.org/">http://www.clustal.org/</a>   |
| DNA CALCULATOR          | <a href="http://www.sigma-genosys.com/calc/DNACalc.asp">http://www.sigma-genosys.com/calc/DNACalc.asp</a>               |
| EXPASY PORTAL           | <a href="http://expasy.org/">http://expasy.org/</a>   |
| FIGTREE 1.3.1           | <a href="http://tree.bio.ed.ac.uk/software/figtree/">http://tree.bio.ed.ac.uk/software/figtree/</a>                     |
| FINCHTV                 | <a href="http://www.geospiza.com/Products/finchtml">http://www.geospiza.com/Products/finchtml</a>                       |
| FR33.NET TRANSLATOR     | <a href="http://www.fr33.net/">http://www.fr33.net/</a>   |
| GENEDOC                 | <a href="http://www.nrbsc.org/gfx/genedoc/">http://www.nrbsc.org/gfx/genedoc/</a>                                       |
| GENEIOUS                | <a href="http://www.geneious.com/">http://www.geneious.com/</a>   |
| MAFFT 6                 | <a href="http://mafft.cbrc.jp/alignment/software/">http://mafft.cbrc.jp/alignment/software/</a>                         |
| MEGA5                   | <a href="http://www.megasoftware.net/">http://www.megasoftware.net/</a>   |
| MUSCLE                  | <a href="http://www.drive5.com/muscle/">http://www.drive5.com/muscle/</a>   |
| NCBI BLAST <sup>®</sup> | <a href="http://blast.ncbi.nlm.nih.gov/">http://blast.ncbi.nlm.nih.gov/</a>   |
| NETNGLYC 1.0 SERVER     | <a href="http://www.cbs.dtu.dk/services/NetNGlyc/">http://www.cbs.dtu.dk/services/NetNGlyc/</a>                         |
| PHYLOGENY.FR            | <a href="http://www.phylogeny.fr/">http://www.phylogeny.fr/</a>   |
| PRIMER III              | <a href="http://biotools.umassmed.edu/bioapps/primer3_www.cgi">http://biotools.umassmed.edu/bioapps/primer3_www.cgi</a> |
| SIGNALP 3.0 SERVER      | <a href="http://www.cbs.dtu.dk/services/SignalP/">http://www.cbs.dtu.dk/services/SignalP/</a>                           |
| TRACER 1.5              | <a href="http://tree.bio.ed.ac.uk/software/tracer/">http://tree.bio.ed.ac.uk/software/tracer/</a>                       |
| TRANSLATORX SERVER      | <a href="http://translatorx.co.uk/">http://translatorx.co.uk/</a>   |
| TREEANNOTATOR 1.6.2     | <a href="http://beast.bio.ed.ac.uk/Main_Page">http://beast.bio.ed.ac.uk/Main_Page</a>                                   |
| TREEILLUSTRATOR         | <a href="http://www.bioinformatics.be/geert/index.php">http://www.bioinformatics.be/geert/index.php</a>                 |
| TREEVIEW                | <a href="http://taxonomy.zoology.gla.ac.uk/rod/treeview.html">http://taxonomy.zoology.gla.ac.uk/rod/treeview.html</a>   |
| TREEVIEW X              | <a href="http://darwin.zoology.gla.ac.uk/~rpage/treeviewx/">http://darwin.zoology.gla.ac.uk/~rpage/treeviewx/</a>       |

## F Primer

### F.1 Vektorspezifische Primer

| Name               | Sequenz (5'-3')        | Länge (bp) |
|--------------------|------------------------|------------|
| M13 <i>forward</i> | GTAAAACGACGGCCAG       | 16         |
| M13 <i>reverse</i> | CAGGAAACAGCTATGAC      | 17         |
| T7 <i>forward</i>  | TAATACGACTCACTATAGGG   | 20         |
| T3 <i>reverse</i>  | AATTAACCCTCACTAAAGGGAA | 22         |

### F.2 Oligo-(dT)-Primer

| Name                | Sequenz (5'-3')   | Länge (bp) |
|---------------------|---|------------|
| Oligo-(dT)-7-Linker | GAGAGAGGATCCAAGTACTAATACGACT<br>CACTATAGGGAGATTTTTTTTTTTTTTTTTT<br>TTTTTTTT | 66         |
| Oligo-(dT)-7-Anker  | GAGAGAGGATCCAAGTACTAATACGACT<br>CACTATAGG                                   | 37         |
| Oligo-(dT)-Linker   | GAGAGAGAGAGAGAGAGAGAAGTAGTC<br>TCGAGTTTTTTTTTTTTTTTTTTTTT                   | 50         |
| Oligo-(dT)-Anker    | GAGAGAAGTAGTCTCGAG  | 18         |

### F.3 Degenerierte Hämocyanin-Primer

| Name                 | Sequenz (5'-3')         | Länge (bp) |
|----------------------|-------------------------|------------|
| PYWDW <i>forward</i> | GGYYTNCNTACTGGGAYTGGAC  | 23         |
| PYWDW <i>reverse</i> | GTGTCCARTCCCAGTANGG     | 19         |
| WAIWQ <i>forward</i> | TGGGCNATRTGGCARGC       | 17         |
| WAIWQ <i>reverse</i> | GCYTGCCAYATNGCCCA       | 17         |
| WHRL <i>reverse</i>  | TGCCARTGNGGRAANACNGCCAT | 23         |

## F.4 Genspezifische Hämocyanin-Primer (MtH<sub>400</sub>-Untereinheit)

| Name                | Sequenz (5'-3')         | Länge (bp) |
|---------------------|-------------------------|------------|
| I_SEQ_forward       | CAAGGTCCTCATCGACTCTG    | 20         |
| II_SEQ_forward      | TGGCAAGTTCACCAGGAT      | 18         |
| II_SEQ_reverse      | TGCAGGAAGTAGTCGTTGTT    | 20         |
| III_SEQ_reverse     | GTCCGTGAATTTCTCGTCAG    | 20         |
| IV_SEQ_reverse      | GTCTCTCCAGGCGTCATAGT    | 20         |
| VI_SEQ_forward      | TACCGAGGACTCGACTACAA    | 20         |
| VII_SEQ_forward     | GTGTTCAACTTCCACCACAA    | 20         |
| A_RACE_forward_1    | CAAGGAACTCTTCAGCAACC    | 20         |
| A_RACE_forward_2    | TCCGATCTGACTGGACG       | 17         |
| A_RACE_reverse_1    | GTCAGCGCTACCGATA        | 16         |
| A_RACE_reverse_2    | GTGAGCGCTATTGGGTTG      | 18         |
| A_RACE_reverse_3    | CTGCAGTCGGAATGGTCA      | 18         |
| A_RACE_reverse_4    | CATGGAGTACTGGTGGCG      | 18         |
| A_forward_2         | TGGAGTGGGACCTGTACC      | 18         |
| A_forward_3         | CCAGTACTCCATGGCCAC      | 18         |
| A_forward_4         | AACATAGACAGGATCTTCGT    | 20         |
| B_RACE_reverse      | CGGCCAAGTCGTTCT         | 15         |
| B_RACE_A1_reverse   | GTGGTGTGGGTCCTGGTATG    | 20         |
| B_RACE_A1_reverse_L | GGGGTGGTGTGGGTCCTGGTATG | 23         |
| B_RACE_A2_reverse   | GGGGATCTGGGCGTTC        | 16         |
| B_RACE_A2_reverse_L | GATGGCGGGGATCTGGGCGTTC  | 22         |
| B_forward_1         | GCCACATACCAGGACCCACA    | 20         |
| B_reverse_1         | GATCTGGGCGTTCGGCAC      | 18         |

|                   |                          |    |
|-------------------|--------------------------|----|
| B_reverse_2       | GGTTCTCGGCCAAGTCGTTC     | 20 |
| B_reverse_3       | ATCTCTGGTGGTTCTGGTGTTTC  | 22 |
| B_reverse_4       | GGAGAATGCCGTCGTACAGTTTAG | 24 |
| B_reverse_5       | CTCGTGCGTCATCGTGTGTG     | 20 |
| C_RACE_A1_reverse | TCCTGGTGGTCTTGTGAC       | 18 |
| C_RACE_A2_reverse | CGACAACCTGCGTGGTG        | 16 |
| C_forward_4       | GCGGTGTAAGCGACAAGAGG     | 20 |
| C_forward_5       | ACACCGTCTTTGCTCTGGAC     | 20 |
| C_reverse_1       | CTGACGAGCAGATCTCGAAG     | 20 |
| C_reverse_2       | TGCGTGATGGCGCAGTTG       | 18 |
| D_forward_2       | GAAGATGGAGGACGACGGAAG    | 21 |
| D_forward_3       | GACGAAGCCCATTACCAAAC     | 20 |
| D_forward_4       | AGGACTCCAGACGTGAACGA     | 20 |
| D_forward_5       | CAATGGAGACCGCAACTG       | 18 |
| D_reverse_1       | CCAGTCGTTTCAGCAGGAGGTT   | 21 |
| D_reverse_2       | GGAAGTGCCGATGTTGTGC      | 19 |
| E_forward_1       | GCGACTTCGAGGTTTCAGTTC    | 20 |
| E_forward_2       | GCCACCGACGCTTTC AAGG     | 19 |
| E_reverse_1       | GGAAGTGGCTCTGGTAATCG     | 20 |
| E_reverse_2       | GGCCTCGTTTCGGGTCTTT      | 18 |
| F_forward_1       | GCAGCCTGTCCGACCTGA       | 18 |
| F_forward_2       | ACAAGATCTGGGCTGTCTGG     | 20 |
| F_forward_3       | CAGAGATCCACAGCCTGAGC     | 20 |
| F_forward_4       | CTCAAGGAGCGTTCCCGC       | 18 |
| F_forward_5       | GAGCTCAAGGAGCGTTCCC      | 19 |
| F_forward_6       | GTATTCAGCCCGAGGACGTG     | 20 |
| F_forward_7       | TGACGCTCTGAAGGACAACG     | 20 |
| F_reverse_1       | CCAGATTTGTCCTTGCAGAT     | 20 |

|             |                       |    |
|-------------|-----------------------|----|
| F_reverse_2 | TCAAAC TCCCAAGGCATTTC | 20 |
| F_reverse_3 | AGGGTTGGAGCATTGATAAC  | 20 |
| G_forward_1 | GGCAAGAAGGTGGCTTGTTG  | 20 |
| G_forward_2 | GTCCACTGAGCTACCAGAAA  | 20 |
| G_forward_3 | TGCCCTGGTCTTTTGATAGG  | 20 |
| G_forward_4 | ACGTCACTICCGTTTTCAGG  | 20 |
| G_reverse_1 | GTGAAGGCCTCAGTCCAGTC  | 20 |
| G_reverse_2 | TGTTGCGGATGATGACGATG  | 20 |
| H_forward_1 | GCAACACTTATGGCAACAACC | 21 |
| H_forward_2 | TTGACGCTCTTGAACACACC  | 20 |
| H_forward_3 | CAGCTCAAGAAGGACAACGA  | 20 |
| H_reverse_1 | GTAGTGGAACCTGGCGGTGT  | 20 |
| H_reverse_2 | CGTTGAACGGCTTCTTTCTC  | 20 |

## F.5 Genspezifische Hämocyanin-Primer (MtH<sub>550</sub>-Untereinheit)

| Name              | Sequenz (5'-3')             | Länge (bp) |
|-------------------|-----------------------------|------------|
| III_SEQ_forward   | GAAGAGGTACTGGAGAAGCA        | 20         |
| VII_SEQ_reverse_1 | GCTGGTTGATGCTCTTG           | 17         |
| VII_SEQ_reverse_2 | GCTCTCCTCATTGAAGAAG         | 19         |
| A_forward_1       | ACATTGTTGTCCGTAAGGATGTGAGC  | 26         |
| A_reverse_1       | GCCAGACGGAGTCAGCCTTGCTATG   | 25         |
| A_reverse_2       | GGCGACGGGTATCTCTGGGGT       | 21         |
| B_RACE_A1_reverse | TCAGCCGCCAGCTTAG            | 16         |
| B_RACE_reverse    | GAGCTTCGTAGGGGAGTC          | 18         |
| B_RACE_reverse_1  | CGTAGGGGAGTCCTCTTC          | 18         |
| B_RACE_reverse_2  | CAGCAGCATACCATCGAAGAGGTCTGT | 27         |
| B_RACE_reverse_3  | GGTAGGTAGCCTCAGCCGCCAGCTTA  | 26         |

|              |                       |    |
|--------------|-----------------------|----|
| B_forward_1  | CAACACCCACATCCCTAAGC  | 20 |
| B_reverse_1  | TAAGCGTCACACCCTCATGG  | 20 |
| B_reverse_2  | CATGGTGAGGGTCCTGGTAG  | 20 |
| C_forward_1  | CAGTGTAGTGCGGCTGGT    | 18 |
| C_forward_2  | GTAGTGCGGCTGGTGAA     | 17 |
| C_reverse_1  | CGCATAGAGGCCAAAGCG    | 18 |
| C_reverse_2  | CATAGAGGCCAAAGCGCAGT  | 20 |
| D_forward_1  | TTTCCAAGGCGACCTACTTC  | 20 |
| D_forward_2  | CGTTTCCAAGGCGACCTAC   | 19 |
| D_reverse_1  | GAATAGTCGAGGGTGGAAAGC | 21 |
| D_reverse_2  | TAGTCGAGGGTGGAAAGC    | 18 |
| E_forward_1  | AGCCCACTGAGAGACAATTC  | 21 |
| E_forward_2  | TACTGTGGATGACGTGGAC   | 19 |
| E_forward_3  | CATGGACAAGCTCAGGAAGG  | 20 |
| E_forward_4  | GCGACTTCGAGGTTTCAGTTC | 20 |
| E_forward_5  | TGGAGTTCTCTGTCTGCAAGG | 21 |
| E_forward_6  | TCATCATCCAGGCTCGTTC   | 19 |
| E_reverse_1  | CCTTAGGACCCTCCTTCCTG  | 20 |
| E_reverse_2  | ACCCTCCTTCCTGAGCTTGT  | 20 |
| F_forward_1  | AGGACGGCAAGCACTCTATCC | 21 |
| F_forward_2  | GCTGGAGTGACGTTTGCT    | 18 |
| F_forward_3  | AGGAGATGCCTTGGTCTTTC  | 20 |
| F_reverse_1  | GGAAGGAGGCAAGGCT      | 16 |
| F_reverse_2  | GTTGTGGGTCACCTCGAACT  | 20 |
| F1_forward_1 | AGACCTTCCACGACCCAGAC  | 20 |
| F1_forward_2 | ACCTTCCACGACCCAGACAC  | 20 |
| F1_forward_3 | CTTTCGACCGTCTCTACAAG  | 20 |
| F1_reverse_1 | GGGAGTGTTTCGTCGTGT    | 17 |

|                          |                           |    |
|--------------------------|---------------------------|----|
| F <sub>1_reverse_2</sub> | CCTTGGTGACTCGCTTC         | 17 |
| F <sub>1_reverse_3</sub> | CGTCCTTGGTGACTCGCTTC      | 20 |
| F <sub>1_reverse_4</sub> | AGCCAGGGAGTGTTCGTCGT      | 20 |
| F <sub>1_reverse_5</sub> | TTCAGGTACTTGGCGGCAAC      | 20 |
| F <sub>1_reverse_6</sub> | GCAACGCAGTCAGCAGAGAC      | 20 |
| F <sub>2_forward_1</sub> | GCGACTTACTTCAACGATAACTC   | 23 |
| F <sub>2_forward_2</sub> | GCGACGTGGAAAGGAG          | 16 |
| F <sub>2_reverse_1</sub> | TGCGACCACCAATCAGGTAG      | 20 |
| F <sub>2_reverse_2</sub> | CCTCCGATCTCCAGGTTGTC      | 20 |
| F <sub>3_forward_1</sub> | TACTGTGCGTGATGTCAG        | 20 |
| F <sub>3_reverse_1</sub> | CCTTGGCAGACATATGGTT       | 19 |
| F <sub>3_reverse_2</sub> | AAACAGCACGTGGTTGG         | 17 |
| F <sub>3_reverse_3</sub> | CACTCCAGTGTGTCATTG        | 19 |
| F <sub>3_reverse_4</sub> | GGGACCTCCTCTGCTTGATG      | 20 |
| F <sub>4_forward_1</sub> | CAGCGAGCGTGTGAGA          | 16 |
| F <sub>4_forward_2</sub> | AAGACTGCTCTCGTGGATCTG     | 21 |
| F <sub>4_forward_3</sub> | GATGAGGGCAAACCTCGGATG     | 20 |
| F <sub>4_forward_4</sub> | TGGACCCTCCCTGTGAC         | 17 |
| F <sub>4_forward_5</sub> | ACCCCTTCGCCAAGAG          | 16 |
| F <sub>4_forward_6</sub> | AGGTGGAGTTTGACGTGTGC      | 20 |
| F <sub>4_forward_7</sub> | AGATACGAGATCACTGACATTCTGG | 25 |
| F <sub>4_reverse_1</sub> | TGGAGGGCAAACCTCGGTCACA    | 20 |
| F <sub>4_reverse_2</sub> | CGGCCGTGGTCGTCGTTAAA      | 20 |
| F <sub>4_reverse_3</sub> | AGCTCTTGGCGAAGGGGTTGGA    | 22 |
| F <sub>4_reverse_4</sub> | CATGCGGCCGTGGTCGTCGTTAAA  | 24 |
| F <sub>4_reverse_5</sub> | TCTTGGAGGTCAGGATGGAG      | 20 |
| F <sub>4_reverse_6</sub> | GTCTTGGAGGTCAGGATGGA      | 20 |
| F <sub>4_reverse_7</sub> | TGGTCGTCGTTAAAGAAGGTCT    | 22 |

|                          |                          |    |
|--------------------------|--------------------------|----|
| F <sub>4_reverse_8</sub> | CCTGCAGTTGCTATCACC       | 18 |
| F <sub>4_reverse_9</sub> | TGCAGCTCGTCCAGAAT        | 17 |
| F <sub>5_forward_1</sub> | GCGCCTCCTTCTTCAATG       | 18 |
| F <sub>5_reverse_1</sub> | GGATCCCTGACGATGTACTG     | 20 |
| F <sub>6_forward_1</sub> | AACGTGCATGACGTCAGCGGTGCT | 24 |
| F <sub>6_forward_2</sub> | ACGAGGCTACCATCCGCGTCAA   | 22 |
| F <sub>6_forward_3</sub> | GGCTAACGAGGCTACCATCC     | 20 |

## G cDNA-Sequenzen

### G.1 MtH<sub>400</sub>-Untereinheit

#### 5'-UTR

AGGCTTCTCCACCTTCAGGGTTCCGCCCTTCTCCTTCTGCAATTTTTCTCTT  
CGGATTGGAGCCCTTCTGAAGGAAGACAAGTTGTGTGCCACGTC

#### Signalpeptid

ATGGTGACGTCACCACGCTTTGCGCTGGCCCTGCTGGCGTTCTGCGTCCAGGTG  
GCCTTTACAACAGGG

#### MtH<sub>400</sub>-a

TCGCTTATTCGCAAGAACGTGGAGAATCTTAGCAGCCAGGAAATCCTCTCCCT  
GCAGACTTCCCTCCGCGCTCTCGCTGCTGATGAAAGTGCTTCTGGATTCAAGGC  
GGTAGCGGCTTACCACGGTGAACCTCCATGTGCGAGGATAACGACGGGGTTG  
CAATTGCCTGCTGTCTTCACGGAATGGCTGTGTTCCACAGTGGCACAGGCTGT  
ACGTCGTGCAGTTCGAGCAGCTGCTGGTCAACAAGGGGCTCAACATCGGTGTT  
CCTTACTGGGACACCCTTGAGCCTCTGACCCACCTGCCGAGCCTTGTGGAGGA  
GCAGATTTTCCGTGACCCTAATGGCGGCGTGGGCGGGAAGAACTCCTGGTTCT  
CTGGTGAGGTCCCCGTGGACGGAGCGATCAAGAACACGGCTCGTGCTGTGGAC  
GACAGGCTGTTTCGAGAAGGTTGCGCCTGGCAAGAACACGCAGCTCTTCAACAT  
CATCCTGGACGCTCTCGAACAGGACGACTACTGCCAGTTCGAGGTCCAGCTCG  
AGGTTGCCACAACCAATCCACTACCTGGTTCGGAGGCCGCCACCAGTACTCC  
ATGGCCACGCTGGAATACAGCTCCTACGATCCCATCTTCTTCCCTCCACCACTCC  
AACATAGACAGGATCTTCGTCATCTGGCAGGAGCTTCAGAAGAAGCGCGGAA  
AGCCCTATGACCATTCGACTGCAGCGTGGAGCAGTTCGCTGGAATTTGAAG  
CCCTTCGACCGCGACACCAACCAATAGCGCTCACCAAGCAGTTCACACGGC  
CAAGGAACTCTTCAGCAACCTGCAGCTCGGCTACGTCTACGACGATCTGACTC  
TCAACGGCATGGACGTCGACCAGCTGTACAACCTCATGGAGCAGCGCCGCGCC  
CGCGAGAGAGCCTTCGGAGTCTTCAGCCTGCACGGTCTGGGTTTCTCCGCCAA

CGCCAGGGTCAAGGTCTGCTCCGATCTGACTGGACGCCGCAAGCGCACTATCG  
GTAGCGCTGACCTCGACGTGGAGCACCCTGCGAGTTCGCCGGCGATTTCTTC  
CTGCTCGGGCGGCCAACGAAATGCCCTGGGAGTTCCACCTGCCCTTCTACTTC  
GACGTCAGCGACGCTGTGCAGCGCCTGGGACTGAGCCTGACCAGCAAGTGGA  
GCCTGGAGTGGGACCTGTACCACGTCAACGGAAGTCTCTGCCCGGAAACCTG  
CTACCTACGCCTTACGGAAAGCACCCGCCCGCCAAGGGTGACCGCGAC

**MtH<sub>400-b</sub>**

TCTGAGGAGGTCAGCTTCTTCGCGGACCGTGACGTGCACAGGCACGCTGTGGT  
GAGGAAGAACGTCGTGGACCTGACTGATGAGGAGGTCTATGAGCTGAGACTG  
GCCATGGATCGCTTCCAGAACGACACATCTGTGGACGGTTACCAGGCCATCGC  
TGAGTTCCACGGCCTGCCTGCTAGATGCCCGAGGCCCGATGCTGCTGTACGTTA  
CGCCTGCTGCATTCACGGCATGGCCACCTTCCCACACTGGCACAGACTGTTTGT  
TGTTTCAGATCGAGGACGAGCTGAAGGCAAGAGACCTGCATTTCCGCATGCCTT  
ACTGGGACTGGACTGTGCCGAACGCCAGATCCCCGCCATCGCTGCTGAAGCC  
ACATACCAGGACCCACACCACCCCGACGTTACCCTGCACAACCCCTTCTACGA  
TGCCGCTGTGGCATTCTGAACACCAGAACCACCAGAGATGTTTCAGAACGACT  
TGGCCGAGAACCCCATCTACGGAGACCACTCTAAACTGTACGACGGCATTCTC  
CTGGCCTTCGAGCAGACCGACTTCTGCGACTTTGAGATCCAGTTCGAGGTCATC  
CACAACCTCCCCCACTTCTGTTGGGCGGCTTCGGGGCCTACACCCTGGCTACC  
CTCCACTACTCGGCTTTCGACCCCATCTTCTACCTGCACCACTCCAACGTCGAC  
AGGATGTGGGCCATCTGGCAGGAGCTGCAGATTCGCCGTGGTCTGCCCTACAG  
GGCTCACTGCGCACACACGATGACGCACGAGCCAATGAAGCCCTTCAGCTTTG  
CTGCTCCCCTCAACAACAACGAGAAGACGTTCTCCCCTCCGAGCCCTCCCAG  
GTCTACGACTACGAGAGGGACCTTGGATACACTTACGACTCCCTCACCTTTGG  
AGGCATGACCATTGAGCAGCTGGGAACTACCTGGAGGAGCGCCAGTCCAAA  
GAGCGCGTCTTTGTGGGCATCGAGCTGCACAACATCGGCGCCTCCGCTTGGGT  
CAAGGTCTCATCGACTCTGGCAGTGGTGGCGAGCCTTACCTGGTCGGCAAGA  
TCGCGGTGCTGGGAGGTGAAAAGGAAA TGCCATGGCACTTCGACCGCCTCTTC  
AAGATGGAGATCACCGATGCTCTGTCTGACCTCGGCCTGAGCTATGACGACAA  
TTTTGACGTCACCCTGCAGATCACTGACGTCACCGGCAAGGAATGGTCCAGCG  
ACTCGTTCAGCATGACACCATCTTCCACGAGCCCGGCAAG

**MtH<sub>400-c</sub>**

ACCTTCGAGGTTTCGCGGTGTAAGCGACAAGAGGACCCGACTGGACGTCGACCA  
CCTCAACAGTGAGCAGATCCAGAACCTTCGCGACGCCTTCAAGGCCCTGGAGG  
ACGACACCTCCATCAACGGCTTTAACCAGATCGCCGCCTTCCACGGACAGCCC  
AGCTGGTGCCTCAGCCCTGATGCTGAGACGAAGTACGCTTGTGCCTGCACGG  
CATGCCTACCTTCCCGCACTGGCACAGGTTGATCACTGTTACGCTGGAGAATGC  
CCTTCGCGACAAAGGTTTTGAGAGTGGTCTCCCCTACTGGGACTGGACGCTGC  
CCATGACCGAGCTTCCCGAGTCCTTACGGCGGAGACGTACGATAATCCCGTC  
AGCGGCGCCAACGAGGCCAACCCCTTACCACGCAGTTGTGCAAGGTCACAA  
GACCACCAGGAGCGTGAGGGGCGAGCTCTTCCAGCAGCCCGCTTTTGGCAAGT  
TCACCAGGATCGGCGAGCAGGTGATGCTGGCCTTTGAGCAGACCGACTACTGT  
TCCTTCGAGGTTTCAGTTCGAGGTGGCCACAACACTACA TCCACGCCCTGGTGGG  
AGGCAACGAGCTGTACTCCATGGCTTCCCTCAGATACACCGCCTACGATCCTCT  
CTTCGCCCTGCACCACTCCAACACTGACAGGATCTGGGCCATCTGGCAGGCTC  
TGCAGAAGTACCGCGGCCTTCCCTACAACACTGCCAACTGCGCCATCACGCAG  
CTGCGCAAGCCCTGCAGCCCTTCGCCCGAGACCAGCACGGTGAACCCTGACTC  
CGTGACACGTGACCACTCCGTGCCTTTCGACGTGTTTCGACTACTCTAGCAGCTT

CCACTATGACTATGACAACCTCGAGTTCAACGGGCTCTCCATCGCCCAGCTGC  
AGCACGAGGTGGTGGTCCGCAAGGCCACGAGCGTGTCTTCGCCGGCTTCATG  
CTGCACGGCATCAAGCAGTCCGCCCTCGTCCGCTTCGAGATCTGCTCGTCAGG  
GGGCGCCTGCGGCCAGGCCGGAGAGTTCTACATCCTGGGAGACGACTATGAGA  
TGCCATGGGAGTATGACAGACTCTTCAAGTACGAGATCACCGACCAGCTGAAG  
GAGCAGGGACTGGAACCTCTTGACCGCTATGACATCAAGTACACCGTCTTTGC  
TCTGGACGGCAGTGTCTGTCGGAAGTACGTTCCGGACAGGTCACCGTTGTGC  
ACACCTATGGTGCT

**MtH<sub>400-d</sub>**

GGTCACGTGGAGCGCAAGTCCACGGCGAGGAGCTGAAGGCTTCTTCCCACGT  
CCGCCGCAACGTTCGATGACCTGACTGGCGGTGAGATCGAGAGCCTGAAGGCTG  
CCCTGCAGAAGATGGAGGACGACGGAAGCTTTGAGAACATCGCCGAGTTCCAC  
GGCGCCCCTGGTCTGTGCGAGCTGAATGGCAAGAAGGTCGGATGCTGCGTGCA  
CGGTATGGCGACTTTCCCGCACTGGCACAGGCTTTACGTTGAGCAGGTGGAGA  
ATGCCCTGCTGTCCCACGGCTCGGCCGTGTCTATACCTACTGGGACTGGACG  
AAGCCCATTACAAACTGCCTGACATCATCGCCCAGGAGACTTACTTCGACTC  
CCGCTCCCAGACCATGCAGAACAAACCCCTTCTTCAGGGGACCCATCAGGACTC  
CAGACGTGAACGACTACACGACCCGCGACCCCCAGCCCGAGCTGTTAAACAAC  
GACTACTTCTGCAGCAGACTCTCCTTGCTTTGGAGCAGACTTCTATTGCGAC  
TTCGAGGCCAGTTTGAGGTGACCCACAACGCTTTCCATTCTTCTGGGTGGA  
CGTGGAAGTACTCCCTCAGTACTCTGGATTACTCCGCCTTCGACCCCGTTTTT  
TTCTGTACCACGCCAACACCCGACAGGATCTTTGCCATCTGGCAGGCTCTCCAG  
AGCTACAGAGGACTGCATTGGGACGAGGCTGACTGCGCTCTGAATCACATGCG  
TGACGGCCTCCATCCGTTTGACGTCAAAGCACAGAACGAATTCGACATCACGC  
ACAAGTACAGCCGCCCCCGTGGATGTCTGGGACTACAGCGACCACCTGCAGTAT  
CACTATGACAACCTCCTGCTGAACGACTGGACCATCCCTCAGCTTGAGGAGGT  
GCTGAAGACCCAGAGGTCACGTGACCGCATGTTCCGCCGGCTTCTCTCTGCACA  
ACATCGGCACTTCCGCTGACGTGGAGCTGCACATCTGTGTGGCTACCGGCAAT  
GGAGACCGCAACTGCAACAACCCCGCTGGCAAGTTCTCCATCTTGGGAGGCGA  
ATATGAAATGCCCTTCACCTTCGACCGTCTTTACAAGTACGACATCAGCGACTC  
CATTAGGAAGCTTGCCCTGAAGCTTGACAGCGGCGCGAACTTTGACCTGAAGG  
TGCAGATCCACGCATACAATGGCTCCTACCTGGACCCAGCCTGCTGAAGGAA  
CCCACCATCATCTTCGAGCCTGGAGAG

**MtH<sub>400-e</sub>**

AACACCTTCCAGACGCAGGACGGCAAGACCCAGCGCAACATGATCAGACGCA  
ACGTCTGACCTCAGCTGGCTGAACGCCGCTCTCTCGTCTTGGCCATGAGAC  
GCCTGCAGGAGGACCACTCCGCCGATGGCTTCCAGGCTCTTGCCCTCTTCCACG  
CCCTGCCAGAACTCTGTCCCTACCCCGAGGCCGCCAAGCGCTACGCCTGCTGC  
GTGCACGGCATGGCCTCTTTCCACAGTGGCACAGACTCTACACCGTGCAGTTT  
GAAGAGGCACTGAGACGCCACGGCTCTCTGGTTCGGCATTCTTACTGGGACAC  
CTCTCGCCCTCAGAAGGCCCTGCCCGCCTTTGCTACTGACGAGAAATTCACGG  
ACCCCGTTCTCAATGTTGAGTTTGACAACCCTTGGCTCGGAGCCGACATCGAAT  
TCGAGAACTCCCACACGGAGAGAGGCCGAACCTTGGCCAGACTGGGAGAGGA  
GGGAGAACACGGATACGACACTTGGCTGTACGAACAGTACCTGCTTGCCCTGG  
AACAGGACAACACTGCGACTTTGAGGTCCAGTTCGAGATCGCCACAACGCC  
ATCCACGCCTGGGTGGGAGGCAGCAAGATTCACAGCATGGGCCACCTTCACTA  
CGCTTCTACGACCCCACTTCTCTCTCCACCCTCTCCACCGACAGGATCTT  
CGCCGTTTGGCAGGAGCTGCAGAGACTCAGGGGCAAAGACCCGAACGAGGCC

AACTGCGCCCTGGAGATCATGCACGAGCCTCTGAAGCCGTTTCAGCTTCGGCTC  
GCCCTACAACCTCAACGAGAACACGCAGCGCTACTCGCGTCCCGAGGACATCT  
TCGATTACCAGAGCCACTTCCACTACGAGTACGACAGCCTGGAGGTTCTCGGC  
ATGAGCCCTGCTCAGCTGGAGGACGCCATTCACAAGAGACAGGAGAAGGACC  
GTGTCTTCGCCGGCTTCTGCTGAAGGGAATCGGCACATCTGCCACGTCGAGT  
TCACCATCTGTAACGAGGCTGGAGTCTGCACAACCGCTGGTGACTTCGACATC  
CTGGGAGGTTCCGCTGAGATGCCCTGGAGCTTTGACCGCGCGTACAAGTACGA  
CATCACC AACGTGCTGGAGAAGAAGGGCCTGGAGGCCACCGACGCTTTCAAG  
GTGAAGGTCGCCATCACGGCCCAGGACGGCACTGCCCTCGATTCCAGTCTCT  
GCCCAGCCCTCCATCATCCTGCAGCACGCCACA

**MtH<sub>400</sub>-f**

CGTGACCAGGGTTCAGTACGAGCTGGCACCCAAACCGTGTCCGCCGCAGCCTGTC  
CGACCTGACGGAGCGCGACGTCA TGAGCCTCAAGTCCGCCCTGCACGACCTTC  
AGGAGGACGACAGCGCCACCGGATGGCAGTCTCTGGCTGCCATACCGGTGTC  
CCCGCCCTCTGCCCGTACCGGAGGAAGCCAAATACGCCCTGCTGTATCCACGG  
CATGCCAACCTTCCCTCACTGGCACCGACTCTACGTTCTGGCCGTGGAGCACTC  
TCTGATCAAGCACGGTAGCTCAGTGGCTATTCCCTACTGGGACTGGACCAACC  
CTCTTGACAAGCTGCCCTGAGCTCTTACCAAGCAGACCTACTATGACGCCTGG  
AGAGACGAGGTTTACGACAACCCCTTCGCTCGCGGTTACGTCAAGGGCGACGA  
GGCTTACACCGTCAGAGATCCACAGCCTGAGCTCAAGGAGCGTTCGCGGACG  
GCAAGCACTCTGTGCTGTTGACCAAGGTCCTGCTTGCCCTGGAGCAGACCGAC  
TATTGCGACTTTGAGGTGCAGTACGAGGTCACCCACAATGCCATCCACTACCT  
GGTGGGCGGGCCCCAGACCTTCTCCCTGTCTCCCTCCACTACTCTTCTTACGA  
CCCCATCTTCTTCGTCCACCACTCCTTCGTCGACAAGATCTGGGCTGTCTGGCA  
GGCTCTGCAGAGGAAGCGTGGCAAACCGTCTGAGCGCGCGGACTGTGCCGTC  
ACTACATGTATGACGACATGCATCCCTTCGACTCCGCCAAGCTCGATCCCGAC  
CCCACCATCCGCTCCACGCCAACCCCTCCACCGTCTTCAACATCTACGACCTC  
GGATACAAGTACGACGACTACCAAGATTGGCGGCAAGAACCTGGACGAACCTGG  
AAGCTCTCATCAAGGACAACCAGGCCACCCAGAGTCTTCGCCGGCTTCCAC  
CTCAAGGGCATCGGATCATCTGCCGACGTCGTCTTCAAGATCTGCAAGGACAA  
ATCTGGCGGCCAGTGCACGCGTGCCGGCAGCATCTTCATCCTCGGTGGCGAAA  
AGGAAATGCCCTGGGAGTTTGACCGTCTGTACAAGTATGACATCACTGACGCT  
CTGAAGGACAACGGTATTCAGCCCAGGACGTGTTTCGACGCTGAGGCGCCCTT  
CTACCTCAAGTACGAGGTCACAGCTGTAAACGGATCCACCCCTGCCAAGAGCG  
TTATCAATGCTCCAACCTCATCTTCGTGCCCGCCACT

**MtH<sub>400</sub>-g**

GGTGCCAGTGTTGATGCCAAGTCCTACGCTGTTGCTGGAATGTCTGTCCGTAAG  
GACATCAACACGCTGAGCTCTGCCGAGACCCAGAGCCTGAAGAACTCGCTGCA  
GAAGGTGATGGACGGCGAGGGTCCACTGAGCTACCAGAAAATCGCTGAGATG  
CACGGCTGGCCCGGTGACTGCGAGCAGAACGGCAAGAAGGTGGCTTGTTGCCA  
CCACGGCATGGCCAGCTTCCCTCAGTGGCACAGGGTCTACACCAGACTTCTGG  
AGATGGCGATGACATGGCAGGGCGCCAACGTGGGTATCCCTACTGGGACTGG  
ACTGAGGCCTTACGGCGCTGCCTTCCCTCGTACGGACGACGGCGACAACCC  
GTTCAACCACGGCCACATCGAGGCTCTCGGCAAGGTGACAAGCCGCGAGCCG  
AGACCACAGCTGTTCAAGGATCCGGAGCACGGCGAGGAGTCCTTCTTCTACCG  
TCAGATCCTCCTGGCCTTCGAGCAGAGAGACTACTGCGATTTTCGAGGTGCAGT  
TTGAGGTCACCCACAATGCCCCCCACTCCTGGATCGGAGGCACCAGCCCCTAC  
GGTATGTCCACCCTGGAATTCTCCTCCTACGACCCCGTCTTCTTCTCCTCACCAC

AGCAACGTCGACAGACAGTTCGCCATCTGGCAGGCCCTGCAGAAGTACCGAG  
GACTCGACTACAACAGCGCCAACTGCGACATCCAGGAGCTGAGGATGGACCT  
GGAGCCCTTCAACCGTGACTACATCCCCATCGTCATCATCCGCAACAACGCC  
GCGCCATCGACGCCTTCAACTACGACCAGTACAGCTACCAGTACGACAACCTC  
AACTTCCACGGCATGTCCATCCCTGAGCTGGAGGCACTCCTGGAGAAGAGGCG  
ATCCGAGGACCATGTCTTCCCTCAACTTCATGCTCCTCGGCATTGGCGCCAGCGC  
TGACGTGACCTTCGATATCTGCGACTCTGAGGGCCACTGCAACTTCGCCGGCA  
CTTTCGCCGTGCTGGGAGGTCCCCTGGAGATGCCCTGGTCTTTTGATAGGCTCT  
TCAAGTATGACGTCACTTCCGTTTTTCAGGCAGATGCACCTCCGCCCCGACTCCG  
ACTACAGCGTCAAGGTCCACATCGTTTCTGTGAATGGTACGGAGCTGGATTCC  
GATCTGCTGGAGGCTCCCTCTGTTAGCTTTGTGCCTGGCAA

### **MtH<sub>400</sub>-h**

CGCGGCCGCAAGGCTCGTGAGGACTACCGCCCGACCGTGAGCGAGGATCTGAT  
CCGTAAAGAGGTGTCTTCCCTGAGCCTGGAGGAGGTCTCCAACCTCAAGAACG  
CCCTCTACAAGCTGCAGAACGACCACGGACCCAACGGCTTCGAAGCCATCGCC  
AGCTTCCACGGCGACCCTGGTCTCTGCCCCGAGGGCTCCGGCTCCAAGTATGC  
TTGCTGCCAGCACGGTATGCCCGTGTTCCCCCACTGGCATCGTCTGCTGACCGT  
CCAGTTTGAACGCGCCCTCAAGGCCAAGGGATCCGTGGTGGGAGTGCCTTACT  
GGGACTGGACGCGCCCCAGCCGTGGCATCCCTCAGCTCTTCAGCAACACTTAT  
GGCAACAACCCCTTCTTGACGTACGAGATCCAGGCGGCCAGCGCCTTCGTTGA  
GAGAGACGTGCAGGTTGACGCTCTTGAACACACCACTGGTCACGAGGGGTTCCG  
AAGATTCCCTCTTCCACCAGGCTTTGGAGACGCTGGAGGAGACGAGCTACTGT  
GAGTTCGAGGTGCAGTTCGAGATGCTGCACAACGCGGTGCACGCACTGGTGGG  
AGGCACGAAGACCCACTCCATGGCCACGCTGGAGTGGTCGGCCTTTGACCCCT  
TCTTCATGGTGCACCACTCCTCCATCGACAGGATCTGGCGCATCTGGCAGGAG  
CTGCAGAGACTGAGAAAGAAGCCGTTCAACGTTGCTCGCTGTGCCATGCGCTA  
CCTGAGGAAACCGCTGGAGCCCTTCTCCTACGCCAGCGTCAACACCGACGAGG  
TCGCCCGCACCAACAGCCGCCCCATCGACATCTTCGACACCGCCAAGTTCCAC  
TACGACTTCGACAACCTTGACCTCGGAGGTCACAGCATCGGCGAGGTCAACAA  
GATGATCAACGACATGAGATCCAAGACCCGCTGTGTTTGTGGCTTTGTGCTGTC  
CGGTATCGAGACAAGTGCCACTGTCAAGATTGAAAATGGACGACAGCCAGGGA  
AATAGCCATGAGGTCGGCACTTTCTACGTGCTGGGAGGACACAACGAGATGCC  
GTGGGCGTACGAGAGAGTGTACAAGTACGAGATCACTGAGGCAGCCAAGAAG  
TACGGACTAGACCACGACTCTGTGTTCAACTTCCACCACAAGTTGTTAAGTAC  
GACGGACAGGAGCTGAGCGCCCAGTTCAAGGCCCCCATCCTGGTGGAGCGCCC  
CGCTGGCGTCTGACTACGATGTGGCCATCTTCCAGCTCAAGAAGGACAACGATC  
TGGCACCCAAGGTTGTTCTGCCAGAGGCACCCGAATTCGATTCCACTCCGCT  
GAGGACGGTGTATCCAGCGTGCTGAGAGAGATGGGCTCCTACACTAACGCTGT  
GTACTGCTCCATCCCGCCCCGGCGACGCCAACACCTACAAACTGGAGGTGGAGT  
ACACCCTGGAGCCTGGTGACTACTTCTTCGTGTCCAATGACGAAGAGAAGTGC  
AAGCAGGGAGACCGCATCCAGATCAGCATCGACGAGGAG

### **3'-UTR**

TAGATGGCGGATGCTGTATCAGACAACAGACTTTATTCAATCTTGGAGACTTTA  
CAAACCATTTGTATAGCCTTGACTTGCCCTTGCCTGTCAGTTGTGCATCTACTCC  
CAGAGTTTCAAAGATTTAACCCTTTACCTCCAGCCGCGGTTATGGCCCATGCC  
GAACGCATAGGCCATAACCCCGGCTGGAAGTGAAAGATTTATCTCTTGACCTC  
CTCGAAACTGTACTCCGCCATGCTGGAAGCTTGGTAGCTACAGATTCTGGGAG  
CACACATATTTACCCTATAGTCACCTGTGGCCCATGCTGAAAGCATGGGCC

ACAGGCGACTATGGTGGTGAAAGGGTGAAGAGGTAGCTTCGCCACATCATAG  
GGTGGAGATACCGAGTCTTTGAAGCTGCTGGAGTAGCAGTGTGGATGGGAGAT  
TTAGATAGTGTGTGAGTGTGGTGGAAACTGCAATAAACTCAGTGCAATACT  
GTAAAAAAAAAAAAAAAAAAAA

## G.2 MtH<sub>550</sub>-Untereinheit

### 5'-UTR

AGTTGATCCCGGGGCTCCACTCAGAGCCTCAGTCCTCACCCGTCTCCGGTCGTG  
TGCTACGCCTGCAAGGTCAAA

### Signalpeptid

ATGACCTGGCTGGCATTTCGCGCTGTTGGTCCTTGGCGCCTTGGTGTCACAATCT  
GCGGCA

### MtH<sub>550</sub>-a

GCACTACTGGTCCGAAAGAAATGTGGAAGACCTGACCCAAGAGGAGGTTCTGTC  
TCTTCAGTCCTACCTCTCCAAACTGGAAGCTGATGAATCGGAATATGGATTTCA  
GAACTAGCGGCATACCACGGAGAACCGTCCATGTGTCAGAGTGACGATGGTA  
CTGCACGTGCTTGCTGTATTACGGCATGGCCAACCTCCCTCAGTGGCACAGAC  
TGTACGTCGTGCAGTTTGAGCAGCTGCTGAGAAAGAAGGGCATGTCCATGGGT  
GTGCCGTACTGGGACACGGTGGGAGAGCTGACTTCTCTGCCTGATCTGGTAGA  
GGAGCACATCTTCCGTGACCCTAACGGCGGGCCGGGGCAAGGAGAACGCCTGG  
TTCTCGGCCAACATCAAGCACGGCAACGTAGACACCAAGACGGCTCGTGCGGT  
CGACAAAAGGCTGTTCCAGAAGGTGGGAGATGGCGAGAAGACCCAGCTCTTC  
AACCTCATCTTGGACGCTCTGGAGAAGGACGACTACTGCCAGTTTGAGGTGCA  
GTTTCGAGGTGGCCACAACCACATCCACTACCTGGTGGGAGGCCGCCACAAAT  
ACTCCAATGTCCACGCTGGAATACAGCTCATAACGACCCCATCTTCTTCCTTACC  
ACTCCGCCATGGACAGGATCTTCGTCACTTGGCAGGAGCTGCAGAAAAAGCGC  
GGTAAGTCTCTGACCATGCCGACTGCAGCGTGGAGCTCTTCCGCTGGAACCT  
GTGGCCCTTTGACCGCTCCAGCAACCCCGTCGAGCTCACCAAGCAGCATAGCA  
AGGCTGACTCCGTCTGGCACCATCCCGAGCTGGGCTACTGGTACGACACTCTG  
ACTCTGGGCGGCATGGACCTTGACCAGCTGCACGCGTACCTGGAGAAGCGCCG  
TGCCAGGGAGAGGGCCTACGCCACCTTCAGTCTCCACGGTGTGGGATACTCTG  
CTAACGCCAGAGTCCAGGTCTGCGGTCTTAGCAGCGAGCACCCAGAGATACC  
CGTCGCCGACGTGATGCTAGTGACAACCTGCGAGTTTGCCGGCGATTTCTTCTG  
CTGGGCGGCCAACGAGATGCCCTGGGAGTTCCACCTGCCCTACTACTTTGA  
CGTCACTGAGGCCGTGCAGAGACTGGGTGTGGCCCTGGATGGCGCCTGGACTG  
TCAAGGTCGACTTTTATCACGTCAACGGCACTGCCATGCCCGGCGACATCCTG  
CCTCCCCCTGTTGGCGCACACAGACCCGCTAAGGGAGCTCACGAC

### MtH<sub>550</sub>-b

AGCGATGATATCCACGGTCATTCCGGCCCTGACATTGTTGTCCGTAAGGATGTG  
AGCGAGCTGACTGCCGAGGAGATGTACGAGCTCCGCCAGGCCATGTCCAAGTT  
CCAGAACGACACCTCTGTGGACGGCTACCAGGCAGTCGCTGAGTTCCACGGCC  
TGCTGCCAGGTGCCCGCGCCCTGACGCTGCTGTCCGTTATGCCCTGCTGCATCC  
ACGGCATGGCTACCTTCCCCCACTGGCACAGGCTGTTTGTGTTTCAGATTGAAG

ATGAGCTGCGCGCTCGCGGTCTCCACATCGGTGTTCCCTTACTGGGACTGGACTC  
GCCCCAACACCCACATCCCTAAGCTGGCGGCTGAGGCTACCTACCAGGACCTT  
CACCATGAGGGTGTGACGCTTAATAACCCCTTCTACGAGGCTGCTGTTGCTTTC  
CTGGGCAAAAAGACCGAGAGAGACGTACAATCTCAGCTGTTCGGAGAACCCAC  
CCTTCGGAGACCACACAGACCTCTTCGATGGTATGCTGCTGGCCTTTGAACAG  
ACTGACTTCTGTGACTTCGAGATCCAGTTCGAGGTGACCCACAATGCCCCCAT  
TTCCTGGTGGGAGGCTTCAAGCCCTACTCGCTGGCCACTCTGCACTACTCGGT  
TTTGACCCCATCTTCTACCTGCATCACTCCAACGTGGACAGGCTGTGGGCCATC  
TGGCAGGAAGTGCAGCTCCGAAGAGGACTCCCCTACGAAGCTCACTGCGCCAA  
CTCCATGACCCATGACCTTCTGAAGCCCTTCGGTTTCGAGGCTCCACTCAACAA  
CAACGAGAAGACCCACTCCCACGCCCGCCCCACCAAGATCTACGACTACGAGC  
TGGAAGTGGACTACACCTACGACAACCTGCAGTTCGGAGGCATGACCATTGAG  
CAGCTGGAGCACTACTTGGAGGAGAGACAGGCCAGAGAGCGAGTCTTTGTCTG  
GTCTGGAGTTCACAACATCGGTGTGTCCGCTTGGGCTGATGTCTACGTTGTTG  
CCGGAGGTCAGAGGTACCGCGCTGGTGGTGTTCGCGTGTGGGCGGCGAGCAG  
GAGATGCCGTGGCACTTCGACAGGCTCTTCAAGATGGAGATCACGGACGCTCT  
GAAGGCGCTTGGCCTGGGCTACAACGACGACTTCACCACTCCCTCAACATCA  
CCGACGTCAACGGCAAACACGTGGATGAGAACACCTTCAGCCACGACACCATC  
TTCCATGATCCTCCCAT

**MtH<sub>50</sub>-c**

GTGCACCACGATGCCGGGGATGCCCGTCTCCCGACAAGGTGAGAAGCAACAT  
CGAGTCGCTGACCCGCGAGCAGGTGCAGAACCCTCCGGGACGCTCTGCAGGCTC  
TGGAGGACGATTCTTCCAACGCTGGCTTCGCCAGATCGCTGCCTTCCACGGC  
GAACCCAGATGGTGGCCGTCACCGACGCTGAGAAGAAGTACGCTTGCTGTCT  
GCACGGCATGGCTGTCTTCCCTCACTGGCACAGACTCATCACCGTCCAGGCTG  
AGAACGCTCTGAGAGCTCACGGCTTCGAGGATGGTCTGCCTTACTGGGACTGG  
ACGCGACCGATGAAGGCCCTGCCCCAGGTGGTGGAGAGCGCCAGTACGTCA  
ACCAGCAGACGGGCCAGACCATGGACAACCCCTTCCACCACGGCATCGCCAAC  
GGACACGCCACGGAGCGCGACGTCAGCCCCAAGCTCTTCCAGCAGCCAGAGTT  
TGGCAAGTTCACCGACATCGCGGAGCTGGTGTATGTTTGCCTTCGAGCAGACCG  
ACTTCTGCGACTTCGAGGTCCAGATGGAGATCGCCCACAACACTACATCCACGCC  
CTGGTGGGCGGCAGCAAGCCCTACTCCATGTCCTCCCTCAGGTACACCACCTA  
CGACCCCATCTTCTTCCCTCCACCACTCCAACACTGACAGAATCTGGGCCATTTG  
GCAGGCACTGCAGAAGTACAGAGGACTGCCATAACAACAGCTAACTGCGCTT  
TGGCCTCTATGCGCAAACCTCTGCAGCCTTTCGCCCAGACGAGCGTGGTGAAC  
CCTGACCCCATGACCCGCGACCATTCGTGCCCTTCGAGATCTTTGACTACCAG  
AGTAGCTTCCACTACACCTACGACGACCTGGAGTTCAACGGTCTCTCCATCCCA  
CAGCTGCAGCATGAGGTGGTAGTCCGTCAGGCGCATGAGAGGGTCTTCGCTGG  
ATTCATGCTTAAGGGAAATCGGACAGTCTGCCCTCGTTACCTTTGAAATCTGCCA  
AGGCAGTGGCCAGTGTAGTGC GGCTGGTGAATTCTACATGCTTGGCGACGAGT  
ACGAGATGCCGTGGGTGTACGACCGCCTGTTCAAGTACGAGATCACCCAGCAG  
CTGGCTGACAAGGGCCTGAAGCCTTAGACTCGTATTACGTGCAATACCAGGT  
GAAGGACCTGCAGAACAACGTGATTGGCACCGACAATTTTGGCAAAGTCACCA  
TTGTCCATACCTACGGCGCT

**MtH<sub>50</sub>-d**

GGTCATCTTGACCGCAAATCCTACGGCCAGGAAGTGAAGGCTTCCAGCCACGT  
CCGCCAGAACGTGGACAGCCTGAGCGAGGGTCAGAAGGAGAGCTTGAAGGCT  
GCTCTGCAGAAGATGGAGAGCGACGGCAGCTTTGAAAAGATCGCTGAGTTCCA

CGGAGCTCCTGGACTGTGCAAGCACGAGGGACGCAATGTGGGCTGCTGCATCC  
ACGGCATGGCTACTTTCCCTCACTGGCACAGACTCTACGTCGAGCAGGTGGAG  
AACGCCCTCCTGACTCACGGCTCTGAGGTCTCCAATCCCCTACTGGGACTGGACC  
TCCCCCATCAAGGAGCTGCCTGACCTCGTTTCCAAGGCGACCTACTTCGACTCC  
CGCTCGCAGACCATGGAGACCAACCCCTTCTTCAAGGCCAGATCGGCAACAC  
TGGCCATTACACTACCAGGGATCCACAGCCCCAGCTGTTC AACACCAAGTACT  
TCCTGAACAATGTCTTCTGGCCCTGGAGCAGACGTCCTTCTGTGACTTTGAGG  
CTCAGTTCGAGGTTGTGCACAACATGATGCACTCTTTTGTGGTGGCCGCGGCA  
AGTACTCGCTTTCCACCCCTCGACTATTCCGCCTACGACCCTCTCTTCTTCATCTA  
CCACGCCAACACCGACAGAATCTGGGCCATCTGGCAGGCACTGCAGAGCCAG  
AGAGGTCTGCCTTGGGACGAGTCTGACTGCGCGCTGAACCACATGCACGACCA  
CCTGCACCCCTTCGACGACAAGGCGCAGAACGAGTTTGACATCACGCACAAAT  
ACAACCACCCCGACGACGTCTGGGACTACAGCGACCACCTCGACTACCACTAC  
GACACGCTCGAGCTCAACGGCTGGTCCATCGCCGAGCTGGAAGAGGTACTGGA  
GAAGCAGCGTTCCCGTGACCGCATGTTCCGCGGTTTCATGCTGCACGACATTG  
GTACCAGCGCTGACGTGGAGATCTTCATCTGTCTGGCCGCGGGCAAGAACCAG  
AGAGACTGCAACCACCTGCCGGCAAGTTCTCCGTTCTTGGTGGAGAGTTTGA  
GATGCCCTTCGCTTTTGACAGACAGTACAAGTTCGACATCAGTGAGGCTCTGA  
CCAAGCTGGGCATCAGACTGGGCAGCAATGCTGACTTTGACCTCAAGGTGACC  
GTCCAGTCATACAATGGCTCTTACCTCAGCTCCGACCTCTTGATGCTCCCACC  
ATCATCTTCTCCCCTGGAAAC

**MtH<sub>50-e</sub>**

CGCAAGGTGCAGTCAGACAGACCTGCCGCCGCCGACGGCGACCTGGTTTCGTCC  
CAACGTCTGGTCCCTCAGCTTCCAGGAGCGCCGCTCTCTTGTTCGGCCCTCAG  
GAGCCTGCAGGAGGACC ACTCCGCCGATGGATTCCAGTCCCTGGCCCTCCTTCC  
ACGCTGTACCCGAAC TTTGCCCTACCCCGAGGCTCCAAGAGGTTTCGCCTGC  
TGCGTGCACGGGTCCCCCGCATTCCACACTGGCACCGTCTGTACTGTTCAG  
TTTGAAGAGGCTCTCAAACGCCATGGCTCCCTCGTTGGTCTACCCTACTGGGAC  
ACTACGGATGCTCAGTCTTTCATCCCCAACTTCTGACGGACGAGAAATTCCAC  
GATACTGTGGATGACGTGGACATTCCCAACCCCTGGCTGGGCGCCGACATCGA  
GTTTGAGAACAGCCACACTGAGAGACAATTCAACATGGACAAGCTCAGGAAG  
GAGGGTCCTAAGGGTTATGACACCTGGTCTTGGGAACAGGTGATGCTGGCCTT  
TGAGCAGGAGAACTTCTGCGACTTCGAGGTTCA GTTCGAGGTGATGCACAACG  
CTATCCATGCCTGGACCGGAGGCACGTCTGTGCACAGCATGGCCCATCTCCAC  
TATGCTTCCCTACGACCCTATCTTCATGCTGCATCACTCTAGCACTGACAGGATC  
TTCGCCATCTGGCAGGAGCTGCAGAGACTGAGAGGCCCTTGACCCCAACTCTGC  
GAACTGTGACCTTGAGACGATGCACGCGCCGCTTAAGCCTTTCAGCTTCGGCTC  
GCCCTACAACCTCAACCCCGACACCAAGCGCTACTCCAGCCCCGACGACATCT  
TCGACTACCACGGACACTTCCACTACGAGTACGACGATCTGGAAATGCAGGGA  
CTTAGCCCACCCAGATCTTGGATCGCA TTGCCAGCGCAAGGAGAGGGAACG  
TGTGTTCCGCGGCTTCTGTCTGCACGGTATCGGTACCTCTGCCACGTGGAGTT  
CTCTGTCTGCAAGGACAATGCCTGCACCCCTGCTGGCGACTTCAACGTCTCGG  
CGGATATGCCGAGATGCCTTGGAGCTTCGACCGTGCGTACAAGTACGAGATCA  
CTGACGTGCTGAAGAAGAACAATCGACCCACGTGACCAGTTCACCATCGAG  
CTGAAGATCACCGCTCTGAACGGCACAACGCTCAGTGGCAACGCCTTTGGCGA  
GCCTTCCATCATCATCCAGGCTCGTTCC

**MtH<sub>550</sub>-f**

GCCGAGCGTGGCCAGTACAACCAGCCAGCCAACCGTATCCGTCACGACCTGAC  
TCACCTGTCAGAGCGCGACTTCATGAGCCTGAAATCCGCCCTCAACGACCTGC  
AGGCTGACACCGGCAAGATGGGATGGCAGAGCCTTGCTCCTTCCACGGCGTC  
CCTGCTCTGTGCCCCGACCCCTGCGGAAGCCACGTACGCTTGCTGTATCCACGGT  
ATGCCACCTTCCCTCACTGGCACAGACTCTACACTCTGCTGGTTGAGCACGCT  
CTGATCGAACACGGCAGTTCTTCGGCTATCCCCTACTGGGACTGGACGCAGCC  
TCTTGATGGACTGCCGGCCCTCTTACCAGCAGACCTACTACGACGCCTGGA  
AGGACAAGGTCTTCGACAACCCCTTCGCTCGCGGATACATCAAGAGCGTTGAC  
GGCTACACGGTGC GCGACCCCCAGTCACAGATGAGTATGCGCTCTAAGGACGG  
CAAGCACTCTATCCTGTTCGACCGCGTGCTCGACGCCCTGGAGCAGGACGACT  
TCTGCGACTTCGAGGTGCAGTTCGAGGTGACCCACAACGCCATCCACTACCTG  
GTCGGCGGTCCACAGACCTACTCCCTCTCATCCCTCCACTACTCTTCCCTACGAC  
CCCATCTTCTTCGTGCACCACAGCTTCGTGCGACAAGATCTGGGCTGTGTGGCAG  
GAGCTGCAGAGGAGACGTGGCAAGCCCTACGACCGTGCTGACTGCGCCGTTAA  
CTACATGAGCAAGCCCATGATGCCTTTCAACTCCCCTGACCTTGACCCTGACTT  
GGAGGTGCGCGCGCATGCTGTGCCTAACTCTGTGTTCAACTACTTTGAGCTGGG  
CTACAAGTACGACAACCTCGAGATCGGCGGCGAGTCGCTGGAGCAGCTGGAG  
GATCGCATCAGGGACCAGCAGAGTCACGCCAGAGTCTTCGCCGGCTTCTACCT  
GGGAGGCATCGGCACCTCGGCTGGAGTGACGTTTGCTGTTTGTAGGAAGGACG  
ACGGCAAGAGGTGCACTTACGGTGGTTCTTTCTTCATCCTGGGAGGATCTAAG  
GAGATGCCTTGGTCTTTCGACCGTGTCTACAAGCACGACATGACGAGGGCCCT  
TCACGCCATTGGCCTCGAGCCTGAGGACGTGTTTCGACTACAAGA ACTACAACG  
TCTTCATTTCTACAACGTCACGGCCGTCAACAACACCGCTCTCCC GCGCTCGA  
TTATTCCCAAGCCTACCATCCTGTATGAACCCGCTAAA

**MtH<sub>550</sub>-f<sub>1</sub>**

GTTCCCGCCAACAGAGTCCGCAAGGACCTTACTGACCTCTCTGCAGAGGACGT  
CAATAGCCTTAAATCCGCTCTGGCTGAGGTT CAGGCCGACAGAGGCAAGCTCG  
GATGGCAGTCTCTGGCCTCCTACCACGGAGTTCCTGCTCTCTGCCCCGACCCCG  
AGGCTGCTGCTTACGCCTGCTGTGTCCACGGTATGCCACCTTCCCTCACTGGC  
ACAGACTCTACACGCTGGAGGTGACCGCGCTCTGATCCACCACGGCAGCAGG  
GTGGCCATCCCTACTGGGACTGGACCCTCCCCATCAAGGAGCTGCCCGCCAT  
CCTGACCGCCAAGACCTTCCACGACCCAGACACCGACGCCGACGTCCCCAACC  
CCTTCGCCCGCGCCTACATCGCTGAGGCCAACACCTACATCCGCAGAGACCCG  
CAGCCCGAGCTGAAGCGAGTCACCAAGGACGGCGAGCACAACGTGCTGTTTG  
ATGAGGTGCTGCTGGCACTCGAGCAGGACGACTTCTGCGACTTCGAGATCCAG  
TTCGAGGTGACCCACAACGCCGTCCACTACCTGGTCCGTGGACACGACGAACA  
CTCCCTGGCTTCCCTCCACTACTCTTCCCTACGACCCCGCCTTCTTCCCTGCACCAC  
AGCTTCGTGCGACAAGATCTGGGCTGTGTGGGAGGAGCTGCAGAAGAAGCGCC  
ACAAGCCTTATGTCTCTGCTGACTGCGTTGCCGCCAAGTACCTGAACGAGCCTC  
TCCACCCCTTCGACGATCCTGACCTTGACCCTGATGACTACGTCCGCACCCACG  
CCATGCCCAGCGACTCCTTCGACTACCACAACCTGGGCTACGACTACGACGAC  
ATGAAGATCGGCGGCAAGAGCCTTGACCAGCTGGCAGACATCATCAAGGAGC  
ACCAGAGCCACTCACGCGCCTTCGCCGGCTTCCACCTGGAGGGCATCGGCGCT  
TCTGCTGACGTGGAGTTCGAGGTGTGCCGCGGGCCAGTGCCAGAAGGCCGG  
AACCTTCTTCTGCTGGGAGGAGCCAAGGAGATGCCGTGGTCTTTCGACCGTCT  
CTACAAGTTTGATATCACCGACCCCTTCCATGCTCTGGGCATCGAGCCTGAGGA  
CGCTATCCTCAGCCCCGCGCCCTTACCTTCAAGATCCAGGTCTTCGCCCTCAA  
CGGCTCCAGCCTGCCTAGCTCCACCCTTGACGCTCCCAAGGTCTTCTACGCCCC

AGCAGCC

**MtH<sub>50</sub>-f<sub>2</sub>**

AATCGTATGCGTCGTAGTGTGCGATGATCTGTCCAATGCCGAGACGACTAGCTT  
GAGAGAGGCCCTTCCCTGAAGATGAGATATGACACTGGTCCTTTGGGCTCCCAGG  
CTTTGGCCTCGTACCACGGTCTGCCCGCTATGTGCCCGACCCCTGCGGAAGCCA  
CGTACGCATGCTGTATCCACGGAATGCCACCTTCCCTCACTGGCACAGACTTT  
ACACTCTGGAGGTTGACCGCTCGCTGGTGAGTCACGGCAGCAGTGTGGCTGTG  
CCCTACTGGGACTGGACTCTGCCTGTCAAGAAGCTGCCTGCCCTGATCGACGA  
CGCGACTTACTTCAACGATAACTCTGGCAAGCAGGAGGAGAACCCCTTCAACT  
GGGGATACATCATGGAGGCCAAGACCAAGACTGTTTCGCGATCCTCAGCCTGAG  
CTCTTCAAGGTGTCCAAGGATGGCCAGCACTCAGTGCTCTTCGACGAGGTTCTC  
CTGGCCCTGGAACAGGACGACTATTGCGACTTCGAGATCCAGTTCGAGGTGAC  
CCACAACGCTCTCCACTACCTGATTGGTGGTTCGCAACCTGTACTCCCTCTCCTC  
TCTCCACTACTCGTCCACGACCCCGCCTTCTTCATGCACCACAGCTTCATCGA  
CAAGATCTGGGTGGTTTGGCAGAAGCTGCAGAAGCGACGTGGAAAGGAGTAC  
CAGACTGCCAAGTGCCTGCCAAGTACATGTCCGAGGCACTCCATCCCTTCGA  
CGATCCTGAGCTGGATCCTGATGAGCTGGTGCCTCAGCCATGCCTGATG  
ACGCTTCAACAACATGGCGCTGGGATACAAGTACGACAACCTGGAGATCGGA  
GGTAAGAACTGGACGAATTGGAGGCTATCATCAAGGAACACCAGAGCCACTC  
TCGCGTGTTCGCCGTTTCCACTTGGAGGGATTCCAGACGTCTGCTGACGTAGT  
CATCAACGTCTGCCGCGATGACACCTGTGCCAGAGCTGGTGCCTTCTTCGTCTC  
GGGAGGAAGCAAGGAGATGCCGTGGTCTTTCGATCGCCAGTTCAGTACGAGA  
TCACCGACGCCCTGGCCGAGCTGAACATCGAGCCTGCCGACGTCACCGACGTC  
AAACGCCCCGCCCTCAGCTTCAAGCTGGACATCCACGCCATCAACGGATCTGC  
CCTGTCCAGCTCCGTTCTGGACGCTCCACCGTCACTCGTCCAGTCTGCCTAC

**MtH<sub>50</sub>-f<sub>3</sub>**

GTTCCCGATGAGCGTGTCCGTCACGACCTGAACCATATGTCTGCCAAGGACCT  
GGTCTCCCTTAGGAACGCCCTCAGAGAGCTCAATGACGACACTGGAGTGCTGG  
GCTTCCAGTCTCTGGCCTCCTTCCATGGTGTGCCCGCCATGTGCCCGACTCCTG  
CTGCTGCCCAGTACGCTTGCTGTATCCACGGTATGCCACCTTCCCTCACTGGC  
ACAGACTCTACACCCTGGAGGTGCAACGTGCTCTGATCTGGCACGGCAGCAAG  
GTGTCCATGCCCTACTGGGACTGGACTCTACCCGTGAGTGAATTGCCCTCCCTG  
GCGACGGAGCAGTCCACTACTACGACCCCAAGAGCGAGACCCGATATGAGAACC  
CCTTCTTAGAGCCTACATTGGCGAAGCTCGTGCCTACACTGTGCGTGATGTCA  
GGCCTGAGTTGACCAAGGTGACGAAGAACGGTGCCAACCACGTGCTGTTTGAT  
GAAGTGATGCTGGCACTGGAACAAGAGGACTACTGTGACTTCGAGATCCAATT  
CGAGGTCACCCACAACGCCATGCACTACGTGGTTCGGCGGCCACCAGCTCCACT  
CCCTCGCCTCCCTTCACTACTCCTCCTACGACCCCATCTTCTTCCCTCCACCCTC  
CTTCGTCGACAAGATCTGGGTCTGCTTGGCAGGAGCTGCAGAAGCGACGCCACA  
AGCCTTACAACACTGCCGAGTGCGCCGCCAAATACCTGGCCGAGCCTCTCCAC  
CCCTTCGATGACCCCGACCTTGACCCTGACGACGTTGTGCGCACCCACGCCAT  
GCCAATGACGTGTTTCGACATCAACGAGCTGGGTACAAGTACGACAACCTTG  
AGATCGGCGGCAAGGACCTGAACCAGCTGGAGGCCATCATCAAGCAGAGGAG  
GTCCACGACAGGGCCTTCGCCGGCTGGAGACTCAAGGGCATCAAGACCTCGG  
CTGACGTCTTCAACGTCTGCACCAAGAACCAGCAGGGCAGCGAGAATTGT  
GTGCGCGGCGGCAAGTTCACTCTGCTGGGTGGAGAGAAGGAGATGCCTTGGGC  
TTTCGACCGTCTTACCTATACGAGATCACGGAGCAGCTTGGTGAGCTGGGCA

TCCCCGTTGAAGAGGCCAGCAAGACCAGCCCCAGCTTCTACTTCACTTGGGAA  
GTTGAGGGAAATCGATGGTTCACATCTGCCCGCCTCCTCCCTCTCCCCGCTACT  
GCCGCTTCAAGGCTGGATAT

**MtH<sub>50</sub>-f<sub>4</sub>**

GACTTCCCCAGCGAGCGTGTGAGACACGAACTGTCTGACCTCTCTGAGAGTGA  
CCTCAGGAGCCTTAAGACTGCTCTCGTGGATCTGCAGGAGGATGAGGGCAAAC  
TCGGATGGCAGTCTCTTGCTGCCTACCACGGAGTCCCCGCCCTCTGCCCCGACCC  
CTGAGGCTGCTACGTACGCCTGCTGTGTCCACGGTATGCCACACCTTCCCTCACT  
GGCACAGACTCTACACGCTGGAGGTGGACCGCGCTCTGCTCCGCCACGGCAGC  
AAGGTGGCCATTCCCTACTGGGACTGGACCCTCCCTGTGACCGAGTTGCCCTCC  
ATCCTGACCTCCAAGACCTTCTTTAACGACGACCACGGCCGCATGGAGTCCAA  
CCCCTTCGCCAAGAGCTACATTAGCGAGTCTAGCAGCTATATCCAGCGCGACC  
CGCAGCCCCGAGCTCTTCAAGAAGTCCAAGGACGGCAAGCATTCCGTGCTCTTC  
GACGAGGTCTGTACGCCCTGGAGCAGCAGGACTACTGCGACTTCGAGGTGCA  
GTTTCGAGGTGACCCACAACGCCTTCCACTACATGCTGGGGCGGTTTCAACGAGC  
ACTCGCTGTCTCGCTGCACTACTCGTCCTACGACCCCGCCTTCTTCTGACCC  
ACTCCTTCGTCGACAAGATCTGGGCTGTGTGGCAGAACTGCAGAAGAAGCGT  
GGCCGCCATAACAACACGGCCCAGTGC GCGGCTGACTACATGGACGAGCCACT  
CCACCCCTTCGACGACCCCGTCTTACCCTGACGAGCTGGTCCGCTCCCACGC  
CCTGCCACCGACGTTTTTACTACCACGACCTCGGCTACACCTATGACGACCT  
CAGCATCGGCGGCCTGAGCCCTTCCAGCTGGCGGACCTGATCGCGAAGGAGG  
AGGGCCACGACAGAGTCTTCGCTGGATTCCAGCTGGCCGGTATTGGCACCTCC  
GCTAAGGTGGAGTTTGACGTGTGCGCTGGCGCCGGTGATAGCAACTGCAGGAG  
GGCTGGTGCCTTCTTCTGCTGGGAGGATCCAAGGAGATGCCGTGGTCTTTTGA  
CCGTCCTTACAGATACGAGATCACTGACATTCTGGACGAGCTGCACATCGAAC  
CTGAGGACGCTACGGACAAGGTCCCGCCCTTCCACTTCAGGACACGCATCACC  
GCTCTTAACGGCTCTGCTCTGTCTTCTGACATCCTTGGCACGCCCAAGATCATC  
TTTGTGGCGTCCAT

**MtH<sub>50</sub>-f<sub>5</sub>**

GTTCCCTCATGACCGCGTCCGTCACGACCTGTCCCGAGTGTCCGAGAGTGACCTT  
CGCAGCCTGAAATCCGCTTTGGGTGAACTGCTCGAGGATGACGGCAAGATGGG  
ATGGCAGGCTCTGGCCTCCTACCACGGAGTCCCCGCCCTCTGCCCCGACCCCTG  
AGGCCGCCACTTACGCCTGCTGTATCCACGGTATGCCACACCTTCCCTCACTGGC  
ACAGACTCTACACGCTGGAGGTGCATCACGCTCTGTACCGCCACGGCAGCAA  
CTGTCGATCCCCTACTGGGACTGGACCCTGCCATGGACAAACTGCCTTCCATC  
CTGACAGCCGACACCTACTACAACGAGAATCACGCCAAACGTGAGGGCAACC  
CCTTCCCTCGGTGCCTACATCAACGAGTCACACCAGTACATCGTCAGGGATCCG  
CAGCCTGCCCTCTGGAAGGTCTCTGGCGATGGTGAACACTCCGTTCTATTTGAA  
GAGGTCTCTATGCTCTGGAACAGGACGACTTCTGCGACTTCGAGGTGCAGTTC  
GAGGTGACCCACAACGCCCTCCACTATCTGGTTCGGCGGCCGTCACCTCTACTC  
GCTGTCTTCGCTGCACTACTCATCCTACGACCCCGCCTTCTTCTGACCCACTC  
CTTCGTCGACAAGATCTGGGCCGTGTGGCAGGAGCTGCAGAAGAGGGCGCGGC  
AAGGCCTACAACCGTGCTGACTGCGCCGCGCGCTTCATGGGGCAGCCTCTCCA  
CCCCTTCGACGACCCCTGACCTTGACCCTGATGAGGAGATCCGCACGCACGCCA  
TGCCCAACGACGTCTTCGATATCCACGAGCTGGGCTACAGCTACGACAACCTTG  
GACATCGGCGGCCATAGCATCGAGGAGCTGGAGCGCCACATCGCCAACCACC  
AGAACCACGACCGCGTCTTCGCTGGATTCCGCCTGTCGGGCATCAAGACCTCG  
GCTGACGTGGAGTTCTTCGTTTGCAAGGGCCCCGACACCACCGGCCCAAGTC

AGATGAGAACTGTGTCCGTGGTGGCGGATTCTTCCTGCTGGGCGGTGAGAAGG  
AGATGCCGTGGTCTTTTCGACCGCCTCTACAAGTACGAGATCACTGACCAGGTG  
AGAGAGCTGGGTATCAATCCTTCGGACGCTGCCGGCGAGGATCCTGGATTCCA  
TTTGACGTCACCTGCCCACGACATCAACGGTGATGAAGTGCCTGCCTCCGCTCT  
GGCCAAGCCTGAGGTGCTCTTCGACGCTGCCCAT

### **MtH<sub>50-f<sub>6</sub></sub>**

CTGGCCCCAGTCCCTCGCGACCGTGTCCGTTACGAACTGTCCAGTGTGTCCGAG  
AGTGACCTCAGCAGCCTGAAGGCCGCCCTGGGTGAACTGCAGCACGACGACG  
GCAAGCTTGGTTGGCAGTCTCTGGCCTCCTACCACGGAGTTCCTGCTCTCTGCC  
CGACCCCCGAGGCTGCCACGTACGCCTGCTGTGTCCACGGTATGCCACCTTCC  
CTCACTGGCACAGACTCTACACGCTGGAGGTTGACCACGCTCTGTTCCGTCAC  
GGCGCCAAGGTGTTCGATCCCCTACTGGGACTGGACGCAGCGCATGGACCGTCT  
GCCCTCCATCCTGACCCGCGCCTCCTTCTTCAATGAGGAGAGCGGCCAGAGCG  
AGCCCAACCCCTTCTCAAGGCCTACATCAATGAGTCACACCAGTACACTGTCT  
CGCACGCCCAATCCGCACTCTGGAAGGTTTCCAGCGACGGCCAGCACTCTGT  
TCTATTCGATGAGGTTCTCTACGCTCTGGAACAGGACGACTTCTGTGACTTCGA  
GGTGCAGTTCGAGGTAACCCACAACGCCATCCACTACCTGGTCGGCGGCCGTA  
ACCTGTACTCGCTGTCTTCGCTGCACTACTCGTCCTACGACCCCGCCTTCTTCT  
GCACCACTCCTTCGTCGACAAGATCTGGGCCGTGTGGCAGGAGCTGCAGAGAC  
GTCGTGGCAAGCCTTACAACACAGCCGACTGTGCGGGCAAGTACCTCAACGAG  
CCGCTCCACCCCTTCGACGACCCTGACCTTGACCCTGATGAGGACGTCCGCGC  
GCACGCCATGCCAGCGACGTCTTTGACTACCGCGGCGAACTTGGCTACTACT  
ATGACAACCTTTAACATCGGCGGCAAGAGCATCAACCAGCTTGAGGCCCTGATC  
CGCCAGCACCGTGGCCACGACCGTGCCTTTGCTGAATTCAGCTTGTGAGGCCA  
AAAGACCTCCGCTGATGTGGAGTCTTTCGCTCTGCAAGAACGACAACCTGTGTGC  
GTGGTGGTGGATTCTTCCTGCTGGGTGGTGAGAAGGAGATGCCTTGGTCTTTTCG  
ACCGTCCCTACAAGTACGAGATCACCGACCAGCTGAAGGAACTGGGTATCGAG  
GCTAACGAGGCTACCAATCCGCGTCAACGTGCATGACGTCAGCGGTGCTGAGCT  
GTCCGCCTCCGTTGTCGGCACGCCAAGAATCACCGTTGACCCCGCCCAC

### **3'-UTR**

TAGTGTAAAAAAAAAAAGTAGTAACCACTAGACAACCTACGCGTGTAGTCAG  
GTCATCATTGTATGTTGTAACAGCTCATTCCATTGGACGGCCTGCTTGCAAAC  
GTGCATTGACAAGACTGATTGGTTGATCCGTTGAGGCCTGAAAGTATATGCTG  
GCCTTGCTGTTAGCATGGCCTGCAGAGACTTGCAAGCCCCAAAGAGTATAAAG  
GACTGAGGTTAGGTAGGGGAAAGGCTTGCAATTGACAGCCCCATTGTCACCTT  
GGTCTAGGTAAAAAAAAAAAAAAAAAAAAA

# H Alignment of the FUs of various Mollusk-Hemocyanins

| Species | 20   | 40   | 60   | 80   | 100  | 120  |  |
|---------|--|--|--|--|--|--|--|
| HtH1_a  | ...DNRV...SL...TVDS...VQALHG...LDVTA...TEPL...SPEDITSY...HAAPAS...DYDKG...--RKT...<br>...ALVRS...DLTSE...ILN...QSLREVND...FENL...FAA...ISY...GPTQ...KDDG...--RDI...<br>...S...L...R...N...V...E...L...S...Q...S...L...R...A...L...A...E...S...E...Y...F...K...L...A...Y...H...E...P...S...M...E...D...N...D...--V...A...I...<br>...L...V...R...N...E...D...L...T...T...A...V...L...A...Q...E...L...K...R...V...Q...E...T...G...P...T...Y...O...A...I...A...Y...H...E...P...G...D...I...C...A...H...D...--H...K...M...<br>...G...N...T...V...R...S...V...D...L...S...A...E...B...E...L...N...I...R...L...K...R...V...Q...E...D...C...I...G...Y...Q...A...L...A...Y...H...E...P...A...G...O...M...E...N...G...--M...K...V...<br>...G...N...L...R...V...S...V...D...L...T...E...E...I...N...L...R...N...L...R...K...V...Q...E...D...E...A...L...F...Q...S...I...G...H...E...P...A...G...O...M...E...N...G...--M...K...V...<br>...P...V...H...R...H...D...D...L...I...R...K...M...D...I...L...T...R...E...E...Y...E...L...Q...V...L...R...A...M...Q...D...E...T...P...T...Y...O...A...I...A...Y...H...E...P...A...D...K...A...P...D...--S...T...V...<br>...P...M...A...H...G...S...E...H...G...V...D...H...G...S...I...R...I...D...I...L...T...R...E...V...V...E...L...R...L...T...K...F...O...N...D...A...S...V...D...Y...O...A...T...E...Y...H...E...P...A...R...P...D...A...K...E...R...V...<br>...S...E...E...V...F...A...D...D...V...H...R...A...V...R...K...M...V...D...L...T...R...E...V...V...E...L...R...L...M...D...F...O...N...D...V...D...Y...O...A...T...E...F...H...E...P...A...R...P...D...A...V...R...Y...<br>...S...D...D...I...H...G...S...G...P...D...I...V...R...K...V...S...E...L...T...A...E...M...Y...E...L...R...Q...M...S...K...F...O...N...D...V...D...Y...O...A...V...E...F...H...E...P...A...R...P...D...A...V...R...Y...<br>...K...P...P...E...D...T...G...L...A...V...R...K...M...D...I...L...T...R...E...E...Y...E...L...Q...V...L...R...A...M...Q...D...E...T...P...T...Y...O...A...I...A...Y...H...E...P...A...D...K...A...P...D...--S...T...V...<br>...A...T...P...H...E...D...P...A...T...L...I...R...K...M...D...I...L...T...R...E...E...Y...E...L...Q...V...L...R...A...M...Q...D...E...T...P...T...Y...O...A...I...A...Y...H...E...P...A...D...K...A...P...D...--S...T...V...<br>...A...A...P...K...G...D...A...D...P...R...K...V...R...N...E...S...L...T...T...E...E...Y...E...L...R...V...M...E...R...F...M...D...K...S...I...D...V...A...L...F...H...E...P...A...R...P...D...A...K...E...R...V...<br>...A...M...V...V...K...T...G...L...N...R...K...M...D...I...L...T...E...E...Y...E...L...R...V...M...E...R...F...M...D...K...S...I...D...V...A...L...F...H...E...P...A...R...P...D...A...K...E...R...V...<br>...P...T...F...E...D...K...H...S...L...R...I...R...K...M...D...I...L...T...P...E...E...T...N...E...L...R...K...L...E...L...L...E...N...I...T...A...G...F...N...Q...L...G...A...F...E...P...K...V...D...N...P...E...A...R...H...K...V...<br>...I...L...P...E...A...Q...I...D...E...V...T...R...K...M...D...I...L...N...A...E...V...L...A...I...R...K...L...A...N...K...E...U...S...I...G...Y...H...E...P...T...L...P...A...D...P...A...E...K...K...V...<br>...T...F...E...Y...R...G...V...D...K...R...T...R...E...V...D...L...N...S...E...Q...I...N...R...D...F...K...A...L...E...D...S...I...N...F...Q...I...A...F...H...E...P...S...M...E...D...N...D...--V...A...I...<br>...V...H...D...A...D...A...D...P...R...K...V...R...N...E...S...L...T...R...E...V...V...E...L...R...L...Q...A...L...E...D...S...I...N...F...Q...I...A...F...H...E...P...S...M...E...D...N...D...--V...A...I...<br>...S...O...L...Y...G...O...E...I...P...T...A...S...R...I...D...L...N...T...G...E...M...E...S...L...R...A...F...L...H...I...Q...D...G...--Y...E...I...Q...Y...H...E...P...G...L...E...N...G...--R...K...V...<br>...G...H...V...E...R...K...S...Y...G...E...L...K...A...S...H...V...R...K...M...D...I...L...T...G...E...E...L...S...K...A...L...Q...M...E...D...G...--S...P...E...N...I...E...F...H...E...P...A...G...L...E...N...G...--K...K...V...<br>...G...H...L...D...R...K...S...Y...G...E...L...K...A...S...H...V...R...K...M...D...I...L...S...E...G...O...K...E...S...L...K...A...L...Q...M...E...D...G...--S...P...E...N...I...E...F...H...E...P...A...G...L...E...N...G...--K...K...V...<br>...G...A...G...V...D...H...E...Y...E...V...T...T...A...S...H...I...R...K...M...D...I...L...N...P...G...E...S...L...K...A...L...Q...M...E...D...G...--S...P...E...N...I...E...F...H...E...P...A...G...L...E...N...G...--K...K...V...<br>...T...Y...G...G...T...E...W...I...P...T...A...S...R...I...D...L...T...L...T...G...E...T...E...S...L...R...N...F...L...R...Q...E...K...--T...K...V...E...Y...I...G...H...E...P...A...R...P...D...A...K...E...R...V...<br>...G...T...Y...G...A...Q...E...W...I...P...T...A...S...R...I...D...L...T...L...T...G...E...T...E...S...L...R...N...F...L...R...Q...E...K...--T...K...V...E...Y...I...G...H...E...P...A...R...P...D...A...K...E...R...V...<br>...G...N...F...G...A...I...W...E...P...T...S...A...N...I...R...K...M...D...I...L...T...D...G...M...E...S...L...R...K...F...D...M...T...G...--Y...E...E...S...F...H...E...P...A...G...L...E...N...G...--K...K...V...<br>...T...H...I...D...H...D...H...E...I...L...V...R...K...M...D...I...L...S...P...R...E...V...S...L...V...K...L...Q...R...K...N...D...R...S...A...D...Y...O...A...I...S...F...H...E...P...A...L...P...E...S...A...A...H...R...Y...<br>...N...E...R...E...A...D...Q...V...G...N...Y...L...V...R...K...M...D...I...L...S...P...O...E...S...L...L...K...M...A...A...L...Q...A...S...A...D...Y...O...A...I...S...F...H...E...P...A...L...P...E...S...A...A...H...R...Y...<br>...N...F...T...Q...D...Q...T...O...R...N...M...I...R...N...V...L...T...S...L...A...E...R...R...S...L...V...L...M...R...R...L...Q...E...D...H...S...A...D...F...O...A...L...S...F...H...E...P...A...L...P...E...P...E...A...K...R...Y...<br>...R...K...V...O...S...D...R...A...A...A...D...G...L...V...R...N...W...S...L...S...F...O...R...E...S...L...V...L...L...R...S...L...Q...E...D...H...S...A...D...F...O...A...L...S...F...H...E...P...A...L...P...E...P...E...A...K...R...Y...<br>...G...A...A...P...H...G...G...S...D...V...T...R...A...E...I...N...T...L...T...P...L...E...V...D...S...L...M...K...M...R...L...Q...E...S...P...N...Y...Q...G...I...S...F...H...A...V...P...L...P...S...P...A...A...N...R...F...<br>...H...S...S...Q...D...H...A...E...D...L...E...V...R...K...M...D...I...L...S...P...R...E...V...S...L...V...K...L...Q...R...K...N...D...R...S...A...D...Y...O...A...I...S...F...H...E...P...A...L...P...E...S...A...A...H...R...Y...<br>...G...S...A...G...N...Q...I...D...P...A...N...R...I...D...L...H...S...E...R...D...F...M...S...L...K...S...A...L...N...D...L...Q...A...T...G...M...Q...L...S...Y...H...G...V...P...A...L...P...E...P...E...A...--A...T...Y...<br>...D...H...S...A...I...K...S...E...G...N...E...Y...L...V...R...N...E...R...L...S...S...M...N...S...L...H...A...L...T...S...L...O...S...S...R...D...I...S...L...K...A...L...T...S...L...O...G...D...Y...O...A...I...S...F...H...E...P...A...L...P...E...S...A...A...H...R...Y...<br>...H...K...L...S...K...H...T...P...N...R...V...E...L...S...S...L...S...R...D...I...S...L...K...A...L...T...S...L...O...S...S...R...D...I...S...L...K...A...L...T...S...L...O...G...D...Y...O...A...I...S...F...H...E...P...A...L...P...E...S...A...A...H...R...Y...<br>...H...D...F...K...R...E...H...V...A...P...N...H...R...K...M...D...I...L...T...L...E...E...R...D...L...Q...S...L...K...A...L...R...D...L...O...H...N...S...N...D...W...A...S...L...S...F...H...S...P...A...K...S...T...P...A...N...E...S...V...<br>...A...D...Q...Y...E...L...A...P...N...R...K...M...D...I...L...T...E...R...D...V...M...S...L...K...S...A...L...H...D...L...Q...E...D...S...A...T...W...Q...S...L...A...Y...G...V...P...A...L...P...E...P...E...A...--A...T...Y...<br>...A...E...R...G...Q...I...D...P...A...N...R...I...D...L...H...S...E...R...D...F...M...S...L...K...S...A...L...N...D...L...Q...A...T...G...M...Q...L...S...Y...H...G...V...P...A...L...P...E...P...E...A...--A...T...Y...<br>...V...P...A...N...R...K...M...D...I...L...T...D...L...S...A...E...D...V...M...S...L...K...S...A...L...E...V...Q...A...R...K...L...M...Q...S...L...S...Y...H...G...V...P...A...L...P...E...P...E...A...--A...T...Y...<br>...N...M...R...S...V...D...L...S...N...A...E...T...S...L...R...E...F...L...K...M...R...Y...T...G...P...L...S...O...A...L...S...Y...H...G...V...P...A...L...P...E...P...E...A...--A...T...Y...<br>...V...P...E...R...V...R...I...L...H...M...S...A...K...D...L...V...S...L...R...N...L...R...E...L...N...D...T...C...V...L...F...O...S...L...S...Y...H...G...V...P...A...L...P...E...P...E...A...--A...T...Y...<br>...D...P...E...S...E...R...V...R...K...M...D...I...L...S...E...D...L...R...S...L...K...T...L...V...D...L...Q...E...D...E...K...L...M...Q...S...L...A...Y...G...V...P...A...L...P...E...P...E...A...--A...T...Y...<br>...V...P...H...D...R...V...R...K...M...D...I...L...T...S...E...D...L...S...L...K...A...L...G...E...L...L...E...D...G...M...Q...L...S...Y...H...G...V...P...A...L...P...E...P...E...A...--A...T...Y...<br>...M...A...P...V...P...R...V...R...K...M...D...I...L...S...E...D...L...S...L...K...A...L...G...E...L...L...E...D...G...M...Q...L...S...Y...H...G...V...P...A...L...P...E...P...E...A...--A...T...Y...<br>...V...D...H...N...D...K...F...A...H...N...Q...V...R...N...I...N...L...D...E...A...D...V...R...S...L...S...L...T...D...L...M...A...E...D...F...O...A...I...R...P...H...G...V...P...O...--G...D...V...<br>...E...S...V...H...D...A...K...S...Y...A...V...A...G...M...S...R...K...M...D...I...L...S...K...R...D...V...M...S...L...K...S...A...L...R...N...L...K...K...K...S...A...D...F...O...S...I...A...F...H...S...P...A...K...S...T...A...D...N...H...P...V...<br>...E...S...V...A...D...A...D...T...P...A...N...R...V...I...S...I...N...E...R...D...I...Q...L...M...A...L...S...E...L...K...E...S...A...D...F...A...A...I...S...F...H...S...P...P...K...T...A...S...G...--N...D...I...<br>...N...E...D...A...D...T...P...L...N...H...I...R...N...V...E...S...D...E...R...I...Q...N...L...M...A...L...T...R...V...K...E...D...S...H...F...O...T...I...S...Y...H...E...T...L...E...P...S...E...R...P...K...Y...<br>...D...H...H...D...H...G...S...I...A...G...V...R...K...M...D...I...L...T...K...A...E...T...D...N...R...E...L...M...G...V...M...A...G...E...N...F...O...A...I...A...F...H...K...P...A...L...P...M...P...D...O...--H...N...S...<br>...G...A...S...E...V...S...S...L...A...G...V...R...K...M...D...I...L...T...S...T...E...D...I...N...R...E...L...R...A...V...Q...A...G...E...N...F...A...S...I...A...F...H...E...P...A...G...E...L...--N...G...--R...R...I...<br>...G...A...S...V...D...A...K...S...Y...A...V...A...G...M...S...R...K...M...D...I...L...S...S...A...T...L...O...K...N...S...L...O...K...V...M...D...E...C...P...L...Y...O...K...I...E...M...W...Q...E...N...G...I...P...Y...M...D...E...A...P...T...A...L...<br>...H...E...S...S...V...G...V...A...H...A...G...R...K...M...D...I...L...T...P...L...S...E...I...D...N...L...K...M...R...A...V...Q...E...D...S...A...T...W...Q...S...L...A...Y...G...V...P...A...L...P...E...P...E...A...--A...T...Y...<br>...G...A...E...H...H...G...D...H...A...G...G...F...G...V...R...K...M...D...I...L...T...P...E...S...I...K...N...R...D...L...R...V...Q...E...D...S...P...H...F...O...A...I...A...Y...H...L...P...P...L...K...S...A...D...O...T...T...L...<br>...T...S...S...R...E...G...D...L...V...G...A...T...I...R...K...M...D...I...L...S...P...E...S...I...E...N...D...L...R...V...Q...E...T...P...H...F...R...A...I...A...Y...H...L...P...A...Q...M...T...P...O...D...E...T...M...<br>...D...F...K...R...E...H...K...T...G...D...I...A...I...R...N...V...L...T...P...S...D...I...K...E...L...R...D...M...A...V...Q...A...T...S...D...N...Y...O...K...I...S...H...Y...E...N...G...--T...A...Y...<br>...H...R...G...K...H...E...D...H...D...R...L...A...D...V...L...I...R...E...V...E...L...S...L...O...B...A...N...I...K...D...L...Y...K...L...O...N...D...S...K...G...F...E...A...I...G...Y...H...E...P...M...E...P...E...R...G...--D...K...V...<br>...K...G...Y...E...R...V...A...A...K...T...A...S...S...L...L...R...A...D...V...D...I...L...T...L...A...E...A...S...N...R...L...D...L...Y...K...L...O...N...D...G...P...N...F...E...A...I...G...Y...H...E...P...A...N...G...--D...K...V...<br>...R...K...G...A...R...E...D...X...R...P...T...S...E...D...L...I...R...E...V...S...L...S...E...A...R...N...L...K...A...L...Y...K...L...O...N...D...G...P...N...F...E...A...I...G...Y...H...E...P...D...P...L...E...P...E...S...G...--S...K...Y...<br>...A...T...H...E...P...H...T...H...E...D...H...R...D...L...R...K...M...D...I...L...S...L...E...A...R...S...L...R...D...L...Y...K...L...O...N...D...G...P...N...F...E...H...I...G...Y...H...E...P...N...P...E...D...S...--S...K...Y...<br>...R...S...S...V...A...Q...T...S...P...P...R...S...V...L...L...E...L...T...L...T...L...E...A...E...N...L...R...E...M...F...O...M...N...E...G...S...S...F...E...H...I...T...F...H...A...P...G...K...P...O...G...A...T...E...N...T...E...S...<br>...C...C...H...G...M...P...S...F...P...H...R...L...Y...V...T...Q...A...E...Q...A...L...Q...A...E...S...N...V...G...I...P...Y...M...D...K...A...M...R...A...L...: 116 | ...C...C...H...G...M...P...S...F...P...H...R...L...Y...V...T...Q...A...E...Q...A...L...Q...A...E...S...N...V...G...I...P...Y...M...D...K...A...M...R...A...L...: 116 | ...C...H...G...M...P...S...F...P...H...R...L...Y...V...T...Q...A...E...Q...A...L...Q...A...E...S...N...V...G...I...P...Y...M...D...K...A...M...R...A...L...: 116 | ...C...H...G...M...P...S...F...P...H...R...L...Y...V...T...Q...A...E...Q...A...L...Q...A...E...S...N...V...G...I...P...Y...M...D...K...A...M...R...A...L...: 116 | ...C...H...G...M...P...S...F...P...H...R...L...Y...V...T...Q...A...E...Q...A...L...Q...A...E...S...N...V...G...I...P...Y...M...D...K...A...M...R...A...L...: 116 | ...C...H...G...M...P...S...F...P...H...R...L...Y...V...T...Q...A...E...Q...A...L...Q...A...E...S...N...V...G...I...P...Y...M...D...K...A...M...R...A...L...: 116 | ...C...H...G...M...P...S...F...P...H...R...L...Y...V...T...Q...A...E...Q...A...L...Q...A...E...S...N...V...G...I...P...Y...M...D...K...A...M...R...A...L...: 116 |



α11

Table with 11 columns labeled α5 to α11. Each column contains a sequence of amino acid residues for various protein entries (Mh1a to Lchh). The sequences are aligned across the columns. A red vertical bar is present on the left side of the table, spanning from the top of the α5 column to the top of the α11 column.

α5

α6

α7

α8

α9

α10

α11



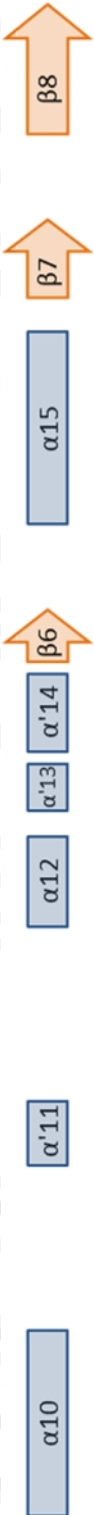
β4

β5

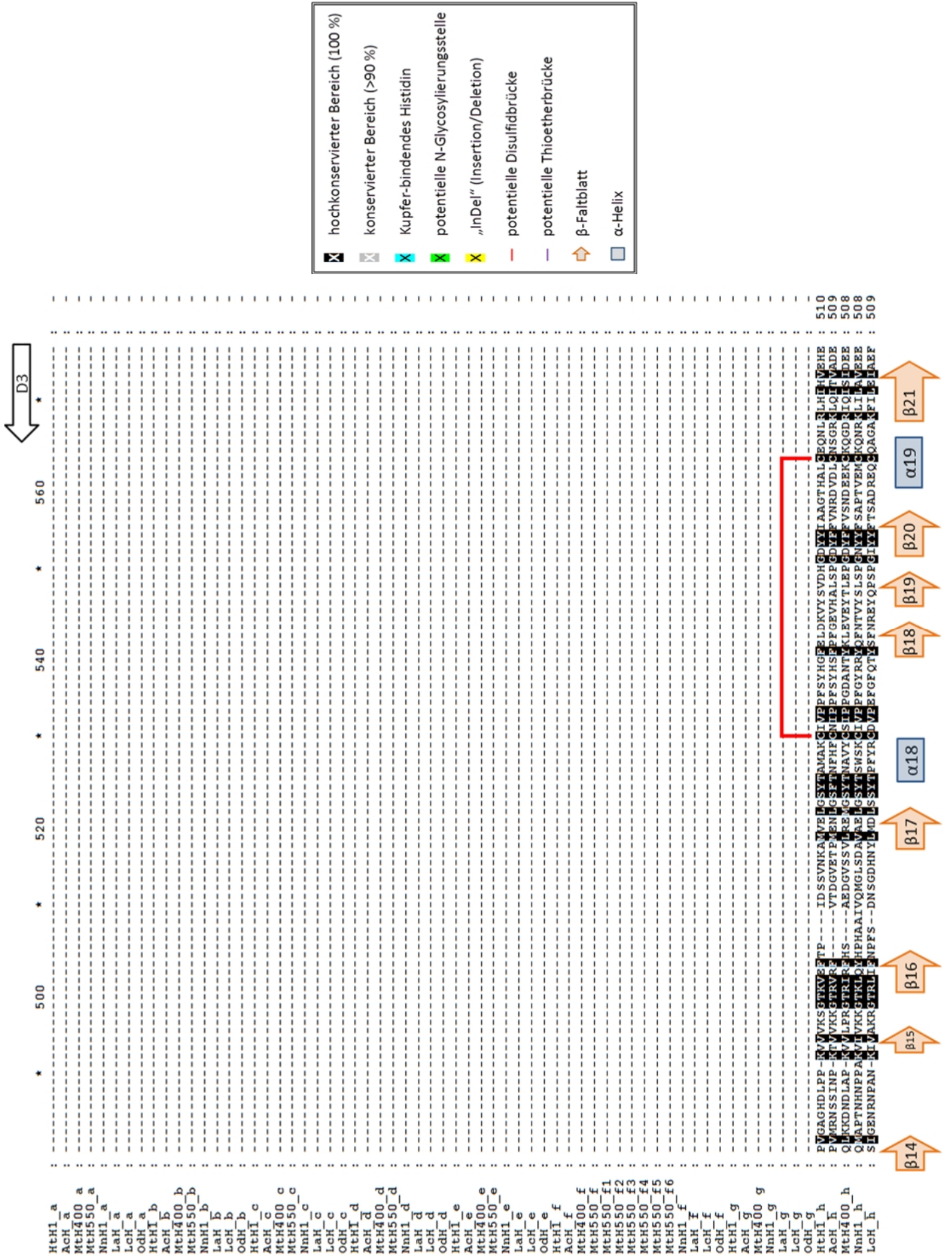
**α8** **β9**

core-domain β-sandwich-domain

|          |        |     |     |     |     |     |     |     |
|----------|--------|-----|-----|-----|-----|-----|-----|-----|
| Hth1_a   | VRLFLV | 260 | 280 | 300 | 320 | 340 | 360 | 324 |
| Ach_a    | VRLVLL | 260 | 280 | 300 | 320 | 340 | 360 | 326 |
| Mth400_a | IRIFV  | 260 | 280 | 300 | 320 | 340 | 360 | 326 |
| Mth50_a  | MRIIF  | 260 | 280 | 300 | 320 | 340 | 360 | 327 |
| Nnh1_a   | VRIIF  | 260 | 280 | 300 | 320 | 340 | 360 | 327 |
| Leh_a    | VRIIF  | 260 | 280 | 300 | 320 | 340 | 360 | 327 |
| Leh_a    | VRIIF  | 260 | 280 | 300 | 320 | 340 | 360 | 325 |
| Odh_a    | VERLFI | 260 | 280 | 300 | 320 | 340 | 360 | 322 |
| Hth1_b   | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 337 |
| Ach_b    | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 338 |
| Mth400_b | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 343 |
| Mth50_b  | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 339 |
| Nnh1_b   | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 337 |
| Leh_b    | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 337 |
| Ach_b    | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 337 |
| Odh_b    | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 337 |
| Hth1_c   | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 337 |
| Ach_c    | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 337 |
| Mth400_c | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 335 |
| Mth50_c  | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 335 |
| Nnh1_c   | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 337 |
| Leh_c    | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 337 |
| Ach_c    | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 337 |
| Odh_c    | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 337 |
| Hth1_d   | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 337 |
| Ach_d    | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 331 |
| Mth400_d | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 333 |
| Mth50_d  | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 331 |
| Nnh1_d   | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 326 |
| Leh_d    | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 326 |
| Ach_d    | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 336 |
| Odh_d    | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 336 |
| Hth1_e   | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 334 |
| Ach_e    | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 341 |
| Mth400_e | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 341 |
| Mth50_e  | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 342 |
| Nnh1_e   | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 340 |
| Leh_e    | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 342 |
| Ach_e    | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 342 |
| Odh_e    | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 344 |
| Hth1_f   | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 334 |
| Ach_f    | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 334 |
| Mth400_f | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 335 |
| Mth50_f  | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 335 |
| Nnh1_f   | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 329 |
| Leh_f    | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 329 |
| Ach_f    | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 325 |
| Mth400_g | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 328 |
| Nnh1_g   | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 329 |
| Leh_g    | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 328 |
| Leh_g    | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 332 |
| Odh_g    | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 331 |
| Hth1_h   | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 334 |
| Ach_h    | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 334 |
| Mth400_h | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 333 |
| Nnh1_h   | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 333 |
| Leh_h    | VRLVLI | 260 | 280 | 300 | 320 | 340 | 360 | 330 |







Der multiple Sequenzvergleich umfasst die FU-a bis FU-h der Hämocyanine von *Haliotis tuberculata* (HtH1), *Aplysia californica* (Ach), *Melanoides tuberculata* (Mth<sub>400</sub>, Mth<sub>550</sub>), *Leptochiton asellus* (LaH), *Lepidochitona cinerea* (LcH) sowie *Octopus dofleini* (OdH). Die Berechnung des Alignments erfolgte mit dem Programm „MAFFT“. Für die grafische Aufbereitung wurde das Programm „GENEDOC“ verwendet. Die Identifikation der  $\alpha$ -Helices (blaue Quadrate) sowie der  $\beta$ -Faltblatt-Strukturen (orangefarbene Pfeile) erfolgte durch den direkten Vergleich mit der OdH-g (Cuff *et al.*, 1998) sowie der KLH1-h (nicht gezeigt; Jaenicke *et al.*, 2010; 2011). Indels der FUs des Mth sind gelb gekennzeichnet. Potentielle N-Glykosylierungsstellen wurden grün markiert. Alle Disulfidbrücken und die Thioetherbrücke wurden rot bzw. violett gefärbt. Die Übergänge zwischen den drei verschiedenen Domänen der funktionellen Einheiten („core-Domäne“, „ $\beta$ -Sandwich-Domäne“, D3 bzw. Tail) wurden oberhalb des Alignments mit weißen Blockpfeilen markiert.

## I Prozentuale Aminosäurezusammensetzung des Mth

|                     | FU                 | FU-a               |                    | FU-b               |                    | FU-c               |                    | FU-d               |                    | FU-e               |                    |
|---------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
|                     |                    | Mth <sub>400</sub> | Mth <sub>550</sub> | Mth <sub>400</sub> | Mth <sub>550</sub> | Mth <sub>400</sub> | Mth <sub>550</sub> | Mth <sub>400</sub> | Mth <sub>550</sub> | Mth <sub>400</sub> | Mth <sub>550</sub> |
|                     | AS                 |                    |                    |                    |                    |                    |                    |                    |                    |                    |                    |
| unpolar / hydrophob | Alanin (A)         | 6,9                | 7,6                | 7,8                | 8                  | 7,5                | 8,5                | 6                  | 6,5                | 9,1                | 7,4                |
|                     | Isoleucin (I)      | 3,5                | 2,1                | 4,8                | 3,9                | 3,6                | 4,1                | 4,6                | 4,1                | 4,3                | 4,3                |
|                     | Leucin (L)         | 11,3               | 9,5                | 7,9                | 8,7                | 8,3                | 7                  | 9,8                | 9,2                | 9,8                | 8,1                |
|                     | Methionin (M)      | 1,4                | 2,1                | 2,1                | 1,9                | 1,5                | 2,7                | 1,4                | 2,2                | 1,7                | 1,9                |
|                     | Phenylalanin (F)   | 5,7                | 5                  | 6,4                | 6,3                | 6,3                | 6                  | 6                  | 6,3                | 5,5                | 6,7                |
|                     | Prolin (P)         | 4,3                | 4                  | 5                  | 5,3                | 4,6                | 5,3                | 4,6                | 3,9                | 4,8                | 6                  |
|                     | Tryptophan (W)     | 1,9                | 2,4                | 1,9                | 1,7                | 1,7                | 1,7                | 1,7                | 1,9                | 1,7                | 2,2                |
|                     | Valin (V)          | 6,6                | 6,6                | 6,4                | 6                  | 6,6                | 6,3                | 4,3                | 4,1                | 4,1                | 4,5                |
|                     | Asparagin (N)      | 5                  | 3,3                | 3,6                | 4,1                | 3,9                | 3,9                | 5,3                | 4,6                | 3,8                | 3,8                |
|                     | Cystein (C)        | 1,7                | 1,7                | 1,2                | 1,2                | 1,7                | 1,7                | 1,7                | 1,7                | 1,7                | 1,7                |
|                     | Glutamin (Q)       | 4,5                | 4                  | 3,6                | 3,9                | 5,3                | 7,2                | 4,1                | 4,6                | 4,5                | 4,1                |
|                     | Glycin (G)         | 7,3                | 7,3                | 5,8                | 6,7                | 6,6                | 6                  | 6,7                | 6,5                | 5,7                | 5,5                |
|                     | Serin (S)          | 6,1                | 6,2                | 5                  | 3,6                | 5,8                | 5,6                | 6,2                | 8                  | 6                  | 7,2                |
|                     | Threonin (T)       | 3,5                | 3,8                | 5,5                | 5,3                | 6,6                | 5,1                | 5                  | 4,6                | 5                  | 4,5                |
| Tyrosin (Y)         | 3,5                | 4,3                | 4,5                | 4,3                | 4,9                | 4,8                | 4,1                | 4,1                | 4,1                | 3,3                |                    |
| basisch             | Arginin (R)        | 5,2                | 5,9                | 4,8                | 4,8                | 4,6                | 3,9                | 4,6                | 3,6                | 5,3                | 5,5                |
|                     | Histidin (H)       | 4,3                | 5,2                | 5,7                | 6,5                | 4,4                | 4,6                | 4,6                | 5,3                | 4,8                | 5                  |
|                     | Lysin (K)          | 4,3                | 5                  | 2,9                | 2,9                | 3,2                | 3,6                | 4,3                | 4,8                | 3,6                | 3,8                |
|                     | Glutaminsäure (E)  | 5,9                | 6,6                | 7,1                | 7,5                | 6,6                | 5,6                | 6,2                | 6,3                | 8,4                | 6,5                |
| sauer               | Asparaginsäure (D) | 7,1                | 7,3                | 8,3                | 7,5                | 6,6                | 6,5                | 8,2                | 8                  | 6,2                | 7,9                |
|                     |                    |                    |                    |                    |                    |                    |                    |                    |                    |                    |                    |
|                     | AS                 |                    |                    |                    |                    |                    |                    |                    |                    |                    |                    |
| unpolar / hydrophob | Alanin (A)         | 7,6                | 7,1                | 10,7               | 8,3                | 7,5                | 7,2                | 8,1                | 7,4                | 6,7                | 6,5                |
|                     | Isoleucin (I)      | 4,3                | 4,3                | 3,9                | 3,9                | 3,6                | 3,6                | 3,6                | 2,9                | 4,2                | 4,1                |
|                     | Leucin (L)         | 8,3                | 7,8                | 9,7                | 9,1                | 8,9                | 10,1               | 9,3                | 8,8                | 7,9                | 7,3                |
|                     | Methionin (M)      | 1,2                | 2,4                | 1                  | 2,5                | 1,9                | 1,2                | 1,4                | 1,2                | 3                  | 2                  |
|                     | Phenylalanin (F)   | 4,5                | 5,7                | 5,4                | 5,4                | 4,8                | 6                  | 5                  | 5,4                | 6,2                | 5,7                |
|                     | Prolin (P)         | 6,2                | 5,9                | 7,1                | 5,4                | 6,3                | 6,3                | 5,7                | 5,7                | 5                  | 4,5                |
|                     | Tryptophan (W)     | 1,9                | 1,9                | 1,7                | 1,7                | 2,2                | 1,7                | 1,9                | 2                  | 2                  | 1,4                |
|                     | Valin (V)          | 6,4                | 5,7                | 5,8                | 6,1                | 6,7                | 5,3                | 5,7                | 6,6                | 6,2                | 7,5                |
|                     | Asparagin (N)      | 2,9                | 3,6                | 2,2                | 3,7                | 3,6                | 1,7                | 3,1                | 3,4                | 4,7                | 4,1                |
|                     | Cystein (C)        | 1,7                | 1,7                | 1,7                | 1,7                | 1,7                | 1,7                | 1,7                | 1,7                | 1,7                | 1,4                |
|                     | Glutamin (Q)       | 3,8                | 3,8                | 2,7                | 3,2                | 3,1                | 2,9                | 2,1                | 3,7                | 4,2                | 3,1                |
|                     | Glycin (G)         | 5,7                | 6,4                | 5,1                | 4,9                | 5,8                | 6                  | 6,9                | 6,4                | 6,7                | 6,7                |
|                     | Serin (S)          | 6,7                | 7,1                | 5,8                | 7,4                | 6,3                | 8,2                | 6,2                | 8,1                | 7,9                | 7,1                |
|                     | Threonin (T)       | 4,3                | 4,5                | 3,9                | 4,7                | 4,6                | 4,6                | 3,3                | 3,7                | 3,7                | 4,7                |
| basisch             | Tyrosin (Y)        | 6                  | 5,7                | 4,1                | 4,4                | 4,8                | 4,3                | 4,5                | 4,9                | 4,2                | 4,3                |
|                     | Arginin (R)        | 4,3                | 5,2                | 3,6                | 4,2                | 4,8                | 4,3                | 4,8                | 5,4                | 4,5                | 5,5                |
|                     | Histidin (H)       | 4                  | 4,3                | 5,6                | 4,2                | 4,8                | 5,1                | 6,7                | 4,9                | 4                  | 4,1                |
|                     | Lysin (K)          | 5,7                | 4,3                | 4,9                | 5,1                | 4,6                | 4,3                | 3,6                | 3,4                | 3,5                | 5,3                |
| sauer               | Glutaminsäure (E)  | 5,5                | 4,8                | 5,6                | 6,1                | 7,5                | 6,3                | 7,4                | 6,6                | 6,2                | 8,3                |
|                     | Asparaginsäure (D) | 9                  | 7,8                | 9,5                | 8,1                | 6,5                | 8,9                | 8,8                | 7,6                | 7,4                | 6,3                |

Die Tabelle zeigt die prozentuale Häufigkeit der 20 proteinogenen Aminosäuren innerhalb der Sequenz der funktionellen Einheiten des Hämocyanins von *Melanoides tuberculata*. Die Aminosäuren wurden anhand ihrer physikalisch-chemischen Eigenschaften in vier Gruppen zusammengefasst: unpolar/hydrophob (Ala, Ile, Leu, Met, Phe, Pro, Trp, Val), polar/neutral (Asp, Cys, Gln, Gly, Ser, Thr, Tyr), basisch (Arg, His, Lys) und sauer (Asp, Gln).

# Literaturverzeichnis

- Abascal , F., Zardoya, R., & Telford, M. J. (2010). TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Research* 38, S. W7-13.
- Albrecht, U., Keller, H., Gebauer, W., & Markl, J. (2001). Rhogocytes (pore cells) as the site of hemocyanin biosynthesis in the marine gastropod *Haliotis tuberculata*. *Cell Tissue Res.* 304, S. 455-462.
- Allen, G. C., Flores-Vergera, M. A., Krasynanski, S., Kumar, S., & Thompson, W. F. (2006). A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. *Nature Protocols* 1, S. 2320 - 2325.
- Altenhein, B. (2000). *Primär- und Genstruktur des Hämocyanins der Schnecke Haliotis tuberculata (Dissertation)*.
- Altenhein, B., Markl, J., & Lieb, B. (2002). Gene structure and hemocyanin isoform HtH2 from the mollusc *Haliotis tuberculata* indicate early and late intron hot spots. *Gene* 301, S. 53-60.
- Babushok , D. V., Ostertag, E. M., & Kazazian, H. H. (2007). Current topics in genome evolution: Molecular mechanisms of new gene formation. *Cell Mol Life Sci* 64, S. 542-554.
- Bailey, J. A., Liu, G., & Eichler, E. E. (2003). An Alu Transposition Model dor the Origin an Expansion of Human Segmental Duplications. *Am J Hum Genet* 73, S. 823-834.
- Bailly, X., Vanin, S., Chabasse, C., Mizuguchi, K., & Vinogradov, S. N. (2008). A phylogenomic profile of hemerythrins, the nonheme diiron binding respiratory proteins. *BMC Evolutionary Biology* 8, 244).
- Bause, E. (1983). Structural requirements of N-glycosylation of Proteins: Studies with proline peptides as conformational probes. *Biochem. J.* 209, S. 331-336.
- Bendtsen, J. D., Nielsen, H., von Heijne, G., & Brunak, S. (2004). Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* 340, S. 783-795.

- Benton, M. J. (1993). *The Fossil Record 2*. London: Chapman and Hall.
- Bergmann, S., Lieb, B., Ruth, P., & Markl, J. (2006). The hemocyanin from a living fossil, the cephalopod *Nautilus pompilius*: Protein structure, gene organization and evolution. *J. Mol. Evol.* 62, S. 362-374.
- Bergmann, S., Markl, J., & Lieb, B. (2007). The first complete cDNA sequence of the hemocyanin from a bivalve, the protobranch *Nucula nucleus*. *J. Mol. Evol.* 64, S. 500-510.
- Bjourson, A. J., & Cooper, J. E. (1992). Band-stab PCR: a simple technique for the purification of individual PCR products. *Nucleic Acids Research*, 20 (17), S. 4675.
- Bowen, N. J., & Jordan, I. K. (2002). Transposable Elements and the Evolution of Eukaryotic Complexity. *Curr Issues Mol Biol* 4, S. 65-76.
- Burmester, T. (2001). Molecular evolution of the arthropod hemocyanin superfamily. *Mol Biol Evol* 18, S. 184-195.
- Burmester, T. (2002). Origin and evolution of arthropod hemocyanins and related proteins. *J Comp Physiol B* 172, S. 95-107.
- Cavener, D. R. (1987). Comparison of the consensus sequences flanking translational start sites in *Drosophila* and vertebrates. *Nucleic Acids Research* 15, S. 1353-1361.
- Chen, J. Y., & Teichert, C. (1983). Cambrian cephalopods. *Geology* 11, S. 647-650.
- Cuff, M. E., Miller, K. I., van Holde, K. E., & Hendrickson, W. A. (1998). Crystal Structure of a Functional Unit from *Octopus* Hemocyanin. *J. Mol. Biol.* 278, S. 855-870.
- Dalbey, R., Lively, M. O., Bron, S., & Maarten Van Dijk, J. (1997). The chemistry and enzymology of the type I signal peptidases. *Protein Science* 6, S. 1129-1138.
- Decker, H., & van Holde, K. E. (2011). *Oxygen and the Evolution of Life*. Berlin Heidelberg: Springer-Verlag.

- Decker, H., Hellmann, N., Jaenicke, E., Lieb, B., Meissner, U., & Markl, J. (2007).  
Minireview: Recent progress in hemocyanin research. *Integr Comp Biol* 47(4), S.  
631-644.
- Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., et al. (2008).  
Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids  
Research* 36, S. 465-469.
- Doyle, J. J., & Doyle, J. L. (1987). A rapid DNA isolation procedure from small quantities  
of fresh leaf tissues. *Phytochem Bull.* 19, S. 11-15.
- Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by  
sampling trees. *BMC Evol. Biol.* 7, 214.
- Drummond, A. J., Ashton, B., Buxton, S., Cheung, M., Cooper, A., Duran, C., et al.  
(2011). Geneious v5.4, Available from <http://www.geneious.com>.
- Edgar, R. C. (2004). MUSCLE: a multiple sequence alignment method with reduced time  
and space complexity. *BMC Bioinformatics* 5(113).
- Feschotte, C., & Pritham, E. J. (2007). DNA Transposons and the Evolution of Eukaryotic  
Genomes. *Annu Rev Genet* 41, S. 331-368.
- Frýda, J., & Blodgett, R. B. (2001). The oldest known heterobranch gastropod,  
Kuskokwimia gen. nov., from the Early Devonian of west-central Alaska, with  
notes on the early phylogeny of higher gastropods. *Bull. Geol. Soc., Prague* 76, S.  
39-53.
- Gatsogiannis, C., & Markl, J. (2009). Keyhole Limpet Hemocyanin: 9-A CryoEM  
Structure and Molecular Model of the KLH1 Didecamer Reveal the Interfaces and  
Intricate Topology of the 160 Functional Units. *J. Mol. Biol.* 385, S. 963-983.
- Gatsogiannis, C., Moeller, A., Depoix, F., Meissner, U., & Markl, J. (2007). *Nautilus  
pompius* Hemocyanin: 9Å Cryo-EM Structure and Molecular Model Reveal the  
Subunit Pathway and the Interfaces between the 70 Functional Units. *J. Mol. Biol*  
374, S. 465-486.

- Gebauer, W., Harris, J. R., Heid, H., Süling, M., Hillenbrand, R., Söhngen, S., et al. (1994). Quarternary structure, subunits and domain patterns of two discrete forms of keyhole limpet hemocyanin: KLH1 and KLH2. *Zoology* 98, S. 51-68.
- Gebauer, W., Stoeva, S., Voelter, W., Dainese, E., Salvato, B., Beltramini, M., et al. (1999). Hemocyanin Subunit Organization of the Gastropod *Rapana thomasiana*. *Arch. Biochem. Biophys.* 372, S. 128-134.
- Goodier, J. L., Ostertag, E. M., & Kazazian, H. H. (2000). Transduction of 3'-flanking sequences is common in L1 retrotransposition. *hum mol gen* 9 (4), S. 653-657.
- Gradstein, F., Ogg, J., & Smith, A. (2004). *A Geologic Time Scale 2004*. Cambridge University Press.
- Hardison, R. C. (1996). A brief history of hemoglobins: plant, animal, protist and bacteria. *Proc Natl Acad Sci USA* 93, S. 5675-5619.
- Harris, J. R., & Markl, J. (2000). Keyhole Limpet Hemocyanin: Molecular Structure of a Potent Marine Immunoactivator. *Eur Urol* 37, S. 24-33.
- Harris, J. R., Gebauer, W., Guderian, F. U., & Markl, J. (1997). Keyhole limpet hemocyanin (KLH): reassociation from Immocuthel<sup>®</sup> followed by separation of KLH1 and KLH2. *Micron* 28, S. 31-41.
- Harris, J. R., Gebauer, W., Söhngen, S. M., Nermut, M. V., & Markl, J. (1997). Keyhole limpet hemocyanin (KLH): characteristic reassociation properties of purified KLH1 and KLH2. *Micron* 28, S. 43-56.
- Herskovitz, T. T. (1988). Recent aspects of the subunit organization and dissociation of hemocyanins. *Comp. Biochem. Physiol. B.* 91, S. 597-611.
- Herskovitz, T. T., & Hamilton, M. G. (1991). Higher order assemblies of molluscan hemocyanins. *Comp. Biochem. Physiol. B.* 99, S. 19-34.
- Hochheimer, A., & Tjian, R. (2003). Diversified transcription initiation complexes expand promotor selectivity and tissue-specific gene expression. *Genes & Development* 17, S. 1309-1320.

- Houbrick, R. S. (1988). Cerithioidean Phylogeny. *Malacological Review* (Supplement 4), S. 88-128.
- Idakieva, K., Severov, S., Svedsen, I., Genov, N., Stoeva, S., Beltramini, M., et al. (1993). Structural Properties of *Rapana thomasiana grosse* Hemocyanin: Isolation, Charakterization, and N-terminal amino acid sequence of two different dissociation products. *Comp. Biochem. Physiol.* 106, S. 53-59.
- Jaenicke, E., & Decker, H. (2004). Functional Changes in the Family of Type 3 Copper Proteins During Evolution. *ChemBioChem* 5, S. 163-169.
- Jaenicke, E., Büchler, K., Decker, H., Markl, J., & Schröder, G. F. (2011). The refined structure of functional unit h of keyhole limpet hemocyanin (KLH1-h) reveals disulfide bridges. *IUBMB Life* 63(3), S. 183-187.
- Jaenicke, E., Büchler, K., Markl, J., Decker, H., & Barends, T. R. (2010). Cupredoxin-like domains in haemocyanins. *Biochem. J.* 426, S. 373-378.
- Katoh, K., Misawa, K., Kuma, K., & Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucl. Acids Res.* 30, S. 3059-3066.
- Kazazian, H. H. (2004). Mobile Elements: Drivers of Genome Evolution. *Science* 303, S. 1626-1632.
- Keller, H., Lieb, B., Altenhein, B., Gebauer, D., Richter, S., Stricker, S., et al. (1999). Abalone (*Haliotis tuberculata*) hemocyanin type 1 (HtH1). Organization of the ~400 kDa subunit, and amino acid sequence of its functional units f, g and h. *Eur. J. Biochem.* 264, S. 27-38.
- Klabunde, T., Eicken, C., Sacchettini, J. C., & Krebs, B. (1998). Crystal structure of a plant catechol oxidase containing a dicopper center. *Nat. Struct. Biol.* 5, S. 1084-1090.
- Kozak, M. (1987). An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleid Acids Research*, 15(20), S. 8125-8148.
- Kump, L. R. (2008). The rise of atmospheric oxygen. *Nature* 451, S. 277-278.

- Kurtz, D. M. (1992). Molecular structure and function relationships of hemerythrins. *Adv Comp Environ Physiol* 13, S. 151-171.
- Lang, W. H., & van Holde, K. E. (1991). Cloning and sequencing of *Octopus dofleini* hemocyanin cDNA: Derived sequences of functional units Ode and Odf. *Proc. Natl. Acad. Sci. USA* 88, S. 244-248.
- Lang, W. H. (1988). cDNA cloning of the *Octopus dofleini* hemocyanin sequence of the carboxy-terminal domain. *Biochem* 27(19), S. 7276-7282.
- Lemon, B., & Tjian, R. (2000). Orchestrated response: a symphony of transcription factors for gene control. *Genes & Development* 14, S. 2551-2569.
- Letunic, I., Copley, R. R., & Bork, P. (2002). Common exon duplication in animals and its role in alternate splicing. *Hum Mol Genet* 11 (13), S. 1561-1567.
- Lieb, B., & Markl, J. (2004). Evolution of molluscan hemocyanins as deduced from DNA sequencing. *Micron* 35, S. 117-119.
- Lieb, B., & Todt, C. (2008). Hemocyanin in mollusks - A molecular survey and new data on hemocyanin genes in Solenogastres and Caudofoveata. *Mol. Phylo. and Evol.* 49, S. 382-385.
- Lieb, B., Altenhein, B., & Markl, J. (2000). The Sequence of a Gastropod Hemocyanin (HtH1 from *Haliotis tuberculata*). *J Biol Chem*, 275(8), S. 5675-5681.
- Lieb, B., Altenhein, B., Lehnert, R., Gebauer, W., & Markl, J. (1999). Subunit organization of the abalone *Haliotis tuberculata* hemocyanin type 2 (HtH2), and the cDNA sequence encoding its functional units d, e, f, g and h. *Eur. J. Biochem.* 265, S. 134-144.
- Lieb, B., Altenhein, B., Markl, J., Vincent, A., van Olden, E., van Holde, K. E., et al. (2001). Structures of two molluscan hemocyanin genes: Significance for gene evolution. *PNAS*, 98(8), S. 4546-4551.
- Lieb, B., Boisguérin, V., Gebauer, W., & Markl, J. (2004). cDNA Sequence, Protein Structure, and Evolution of the Single Hemocyanin from *Aplysia californica*, an Opisthobranch Gastropod. *J. Mol. Evol.* 59, S. 536-545.

- Lieb, B., Gebauer, W., Gatsogiannis, C., Depoix, F., Hellmann, N., Harasewych, M. G., et al. (2010). Molluscan mega-hemocyanin: an ancient oxygen carrier tuned by a ~550kDa polypeptide. *Frontiers in Zoology* 7, S. 14.
- Mankad, R. V., Gimelbrant, A. A., & McClintock, T. S. (1998). Consensus Translational Initiation Sites of Marine Invertebrate Phyla. *Biol. Bull.* 195, S. 251-254.
- Markl, J. (1996). Blaues Blut - Struktur, Funktion und Evolution der Hämocyane. *Chemie in unserer Zeit* 1, S. 6-18.
- Markl, J., & Decker, H. (1992). Molecular structure of the arthropod hemocyanins. In e. Mangum CP, *Physiol, Adv Comp Environ* (S. 325-376). Heidelberg: Springer.
- Markl, J., Lieb, B., Gebauer, W., Altenhein, B., Meissner, U., & Harris, J. R. (2001). Marine tumor vaccine carriers: structure of the molluscan hemocyanins KLH and HtH. *J. Cancer Res. Clin. Oncol.* 127, S. R3-R9.
- Markl, J., Savel-Niemann, A., Wegener-Strake, A., Süling, M., Schneider, A., Gebauer, W., et al. (1991). The Role of Two Distinct Subunit Types in the Architecture of Keyhole Limpet Hemocyanin (KLH). *Naturwissenschaften* 78, S. 512-514.
- Martin, A. G., Depoix, F., Stohr, M., Meissner, U., Hagner-Holler, S., Hammouti, K., et al. (2007). *Limulus polyphemus* Hemocyanin: 10Å Cryo-EM Structure, Sequence Analysis, Molecular Modelling and Rigid-body Fitting Reveal the Interfaces Between the Eight Hexamers. *J. Mol. Biol.* 366, S. 1332-1350.
- Meissner, U., Gatsogiannis, C., Moeller, A., Depoix, F., Harris, J. R., & Markl, J. (2007). Comparative 11 Å structure of two molluscan hemocyanins from 3D cryo-electron microscopy. *Micron* 38, S. 754-765.
- Meissner, U., Dube, P., Harris, J. R., Stark, H., & Markl, J. (2000). Structure of a Molluscan Hemocyanin Didecamer (HtH1 from *Haliotis tuberculata*) at a 12Å Resolution by Cryoelectron Microscopy. *J. Mol. Biol.* 298, S. 21-34.
- Meyer, A., & Lieb, B. (2010). Respiratory proteins in *Sipunculus nudus* - Implications for phylogenie and evolution of the hemerythrin family. *Comp. Biochem. and Physiol.* 155, S. 171-177.

- Miller, K. I., Cuff, M. E., Lang, W. F., Varga-Weisz, P., Field, K. G., & van Holde, K. E. (1998). Sequence of the *Octopus dofleini* Hemocyanin Subunit: Structural and Evolutionary Implications. *J. Mol. Biol.* 278, S. 827-842.
- Möller, V. M. (2008). *Molekularphylogenetische Untersuchungen am Hämocyanin der Käferschnecken (Polyplacophora)*. (Diplomarbeit). Mainz.
- Moran, J. V., DeBerardinis, R. J., & Kazazian, H. H. (1999). Exon shuffling by L1 retrotransposition. *Science* 283, S. 1530-1534.
- Morris, S. C. (2006). Darwin's dilemma: the realities of the Cambrian 'explosion'. *Phil. Trans. R. Soc. B.* 361, S. 1069-1083.
- Neafsey, D. E., & Galagan, J. E. (2007). Positive selection for unpreferred codon usage in eukaryotic genomes. *BMC Evolutionary Biology*, 7(119).
- Nishiguchi, M. K., & Mapes, R. H. (2008). Cephalopoda. In W. F. Ponder, & D. R. Lindberg, *Phylogeny and Evolution of the Mollusca* (S. 163-199). London: University of California Press.
- Ohno, S. (1970). *Evolution by Gene Duplication*. Berlin: Springer.
- Ostertag, E. M., & Kazazian, H. H. (2001). Biology of Mammalian L1 Retrotransposons. *Annu Rev Genet* 35, S. 501-538.
- Perbandt, M., Gutlöhrllein, E. W., Rypniewski, W., Idakiewa, K., Stoeva, S., Voelter, W., et al. (2003). The Structure of a Functional Unit from the Wall of a Gastropod Hemocyanin Offers a Possible Mechanism for Cooperativity. *Biochemistry* 42, S. 6341-6346.
- Raff, R. A., & Raff, E. C. (1970). Respiratory Mechanisms and the Metazoan Fossil Record. *Nature* 228, S. 1003-1005.
- Rambaut, A., & Drummond, A. J. (2007). Tracer v1.5, Available from <http://beast.bio.ed.ac.uk/Tracer>.

- Robinson, M., Lilley, R., Little, S., Emtage, J. S., Yarranton, G., Stephens, P., et al. (1984). Codon usage can affect efficiency of translation of genes in *Echerichia coli*. *Nucleic Acids Research*, 12(17), S. 6663-6671.
- Saitou, N., & Nei, M. (1987). The Neighbor-joining Method: A new Method for Reconstructing Phylogenetic Trees. *Molecular Biology and Evolution*, 4(4), S. 406-425.
- Senozan, N. M., Landrum, J., Bonaventura, J., & Bonaventura, C. (1981). Hemocyanin of the giant keyhole limpet, *Megathura crenulata*. In J. Lamy, & J. Lamy, *Invertebrate oxygen binding proteins* (S. 703-717). New York: Dekker.
- Sessions, A. L., Doughty, D. M., Welander, P. V., Summons, R. E., & Newman, D. K. (2009). The Continuing Puzzle of the Great Oxidation Event. *Curr. Biol.* 19, S. R567-R574.
- Shakin-Eshleman, S. H., Spitalnik, S. L., & Kasturi, L. (1996). The Amino Acid at the X Position of an Asn-X-Ser Sequon Is an Important Determinant of N-Linked Core-glycosylation Efficiency. *J. Biol. Chem.* 271(11), S. 6363-6366.
- Siezen, R. J., & van Bruggen, E. F. (1974). Structure and properties of hemocyanins. XII. Electron microscopy of dissociation products of *Helix pomatia* alpha-hemocyanin: quarternary structure. *J Mol Biol* 90, S. 77-89.
- Silvestro, D., & Michalak, I. (2011). raxmlGUI: a graphical front-end for RAxML. *Org Divers Evol*(DOI 10.1007/s13127-011-0056-0), S. 1-3.
- Söhngen, S. M., Stahlmann, A., Harris, J. R., Müller, S. A., Engel, A., & Markl, J. (1997). Mass determination, subunit organization and control of oligomerization states of keyhole limpet hemocyanin (KLH). *Eur. J. Biochem.* 248, S. 602-614.
- Stagl, V. (1993). Die Bruttasche von *Melanooides tuberculata* (O.F. Müller) (Gastropoda: Thiaridae). *Ann. Naturhist. Mus. Wien* (94/95), S. 187-192.
- Stamatakis, A. (2006a). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22, S. 2688-2690.

- Starmühlner, F. (1969). Ergebnisse der Österreichischen Madagascar-Expedition. *Malacologia*, 8(1-2), S. 1-434.
- Stoeva, S., Schutz, J., Gebauer, W., Hundsdorfer, T., Manz, C., Markl, J., et al. (1999). Primary structure and unusual moiety of functional unit 2-c of keyhole limpet hemocyanin (KLH). *Biochem Biophys Acta*(1435), S. 94-109.
- Streit, K. S. (2008). *Differentielle Expression und molekulare Evolution von Mollusken-Hämocyaninen (Dissertation)*. Mainz.
- Streit, K., Lieb, B., & Markl, J. (2003). Keyhole limpet hemocyanin: genes, cDNA and evolution. *95th Annual Meeting of the DZG, Berlin, Germany*.
- Strong, E. E., Colgan, D. J., Healy, J. M., Lydeard, C., Ponder, W. F., & Glaubrecht, M. (2011). Phylogeny of the gastropod superfamily Cerithioidea using morphology and molecules. *Zoological Journal of the Linnean Society* 162, S. 43-89.
- Swerdlow, R. D., Ebert, R. F., Lee, P., Bonaventura, C., & Miller, K. I. (1996). Keyhole limpet hemocyanin: structural and functional characterization of two different subunits and multimers. *Camp. Biochem. Physiol. B* 113, S. 537-548.
- Szczesna-Skorupa, E., Mead, D. A., & Kemper, B. (1986). The selection of the first AUG as the initiator of eukaryotic mRNAs translation is favored by a 5'-terminal CAP group and a purine in the -3' position. *Biochem. and Biophys. research communications*, 140(1), S. 288-293.
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Kumar, S. (2011). MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution*.
- Terwilliger, N. B. (1998). Functional adaptations of oxygen-transport proteins. *J Exp Biol* 201, S. 1085-1098.
- Tian, B., Hu, J., Zhang, H., & Lutz, C. S. (2005). A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Research*, 33(1), S. 201-212.

- van Holde, K. E., & Miller, K. I. (1995). Hemocyanins. *Adv. Protein Chem.* 47, S. 1-81.
- van Holde, K. E., Miller, K. I., & Decker, H. (2001). Hemocyanins and Invertebrate Evolution. *J. Biol. Chem.* 276, S. 15563-15566.
- Venkataraman, K., Brown, K. M., & Gilmartin, G. M. (2005). Analysis of a noncanonical poly(A) site reveals a tripartite mechanism for vertebrate poly(A) site recognition. *Genes & Development* 19, S. 1315-1327.
- Wahle, E., & Keller, W. (1992). The Biochemistry of 3'-end cleavage and polyadenylation of messenger RNA precursors. *Annu. Rev. Biochem.* 61, S. 419-440.
- Warnke, K. M., Meyer, A., Ebener, B., & Lieb, B. (2011). Assessing divergence time of Spirulida and Sepiida (Cephalopoda) based on hemocyanin sequences. *Mol Phyl and Evol* 58, S. 390-394.
- Weber, R. E., & Vinogradov, S. N. (2001). Nonvertebrate hemoglobins: functions and molecular adaptations. *Physiol Rev* 81, S. 569-628.
- Zelnick, C. R., Burks, D. J., & Duncan, C. H. (1987). A composite transposon 3' to the cow fetal globin gene binds a sequence specific factor. *Nucleic Acids Res* 15(24), S. 10437-10453.
- Zeugin, J. A., & Hartley, J. L. (1985). Ethanol Precipitation of DNA. *focus*, 7(4), S. 1-2.
- Zhang, J. (2003). Evolution by gene duplication: an update. *TRENDS in Ecology and Evolution* 18(6), S. 292-298.

## Danksagung

Herrn [REDACTED] danke ich herzlich für die Aufnahme in seine Arbeitsgruppe sowie die Bereitstellung des spannenden Themas und für die ausgezeichneten Arbeitsbedingungen. Vielen Dank zudem für die Betreuung und Korrektur meiner Arbeit.

Herrn [REDACTED] danke ich vielmals für die Übernahme der Zweitkorrektur.

Bei [REDACTED] bedanke ich mich für die kompetente und freundschaftliche Unterstützung bei meinen kleinen und großen Problemchen im Labor und am Computer. Außerdem nochmals vielen Dank für das Korrekturlesen des Manuskripts und die zahlreichen wertvollen Anregungen und Diskussionen. Muchas gracias por todo y saludos!

[REDACTED] und [REDACTED] danke ich für jegliche Hilfe und Unterstützung im alltäglichen Laborwahnsinn.

Ein ganz besonderes Dankeschön geht an meine aktuellen „Mitbewohner“ [REDACTED] und [REDACTED] sowie die ehemalige „Besatzung“ von Raum 02-262 [REDACTED] und [REDACTED]. Die schöne Zeit mit euch werde ich nie vergessen!

[REDACTED] und natürlich auch allen anderen aktuellen und ehemaligen Mitstreitern danke ich sehr für das hervorragende Arbeitsklima und die tolle Zeit in der [REDACTED].

[REDACTED] und [REDACTED], euch danke ich für eure große Hilfsbereitschaft bei sämtlichen verwaltungstechnischen und organisatorischen Problemen.

[REDACTED], vielen Dank, dass Du in den letzten Jahren immer für mich da warst und an mich geglaubt hast.

Von ganzem Herzen danke ich auch meinen Eltern für ihre liebevolle Unterstützung jeglicher Art und für das Vertrauen, das sie in mich gesetzt haben. Ein ganz großes Dankeschön auch an meine beiden Schwestern [REDACTED] und [REDACTED] – ich bin froh dass es euch gibt! [REDACTED], Dir danke ich für einfach alles, was Du für mich getan hast. Ich weiß, wie gerne Du den Abschluss meiner Arbeit miterlebt hättest.

# CURRICULUM VITAE



## Eidesstattliche Erklärung

Ich erkläre, dass ich die vorgelegte Arbeit selbständig, ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt habe, die ich in der Arbeit angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder nicht veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Bei den von mir durchgeführten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der Satzung der Johannes Gutenberg -Universität Mainz zur Sicherung guter wissenschaftlicher Praxis niedergelegt sind, eingehalten.

---

Datum, Ort

---

Unterschrift

