

Diss. ETH No. 8922

**STRUCTURE AND REGULATION OF THE MELANIN OPERON
FROM *STREPTOMYCES GLAUCESCENS***

A dissertation submitted to the

SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZUERICH
for the degree of
Doctor of Natural Sciences

presented by
MARTIN GEISTLICH
dipl. Natw. ETH
born July 26, 1960
citizen of Schlieren ZH

accepted on the recommendation of
Prof. Dr. R. Hütter, examiner
Prof. Dr. H. Hennecke, co-examiner
PD Dr. K. Lerch, co-examiner

*Eingesehen,
Kontrolliert*

6. August 1989

R. Hütter

Zürich 1989

I. SUMMARY

The melanin operon from *Streptomyces glaucescens* consists of three open reading frames, which are transcribed as one polycistronic mRNA. The second open reading frame corresponds to the structural gene of the enzyme tyrosinase (*melC*). A 402 bp open reading frame (ORF I) exhibiting a presumptive leader peptide sequence precedes the *melC* gene. The gene product of ORF I has been discussed to be involved in copper transport as well as in secretion of tyrosinase.

The nucleotide sequence of a 1489 bp DNA fragment situated downstream of the tyrosinase coding region and containing the third open reading frame (ORF III) of the *mel* operon was determined. The ORF III showed a strongly biased use of codons with G or C in the third position which is characteristic of *Streptomyces* genes, and has the potential to code for a protein of 21'797 daltons. No evidence, however, could be acquired to demonstrate the presence of an ORF III-encoded gene product.

The transcription initiation site of the *mel* operon, determined by S1 mapping and primer elongation experiments, lies 32 to 34 bp upstream of the translation initiation codon of the first open reading frame.

The termination site of the *mel* transcript was determined by nuclease S1 mapping and lies 2056 or 2057 bp downstream of the first transcribed nucleotide (+1) of the operon. Transcription termination occurred at the end of the stem of a potential hairpin-loop structure with a free energy of formation (ΔG^0) of -29kcal implying a termination mechanism comparable to rho-independent termination in *E. coli*.

A total of 172 to 219 bp upstream of the transcription start point are necessary for a fully active and regulated *mel* promoter. Deletion analysis, gel retardation assays and DNase I footprint experiments permitted division of the promoter into three functional domains, which include the RNA polymerase recognition site up to nucleotides -33 to -42, the binding region of a protein of assumed regulatory function between nucleotides -65 and -93, and an upstream activator site, located between positions -158 and -219.

The 402 bp ORF I preceding the *melC* gene was shown to influence expression of the tyrosinase gene. Deletion and nonsense mutations in the ORF I sequence caused a dramatic drop or complete loss of tyrosinase activity. The effect of the deletion mutations could be complemented in *trans* by an intact copy of ORF I, whereas the nonsense mutations withstood complementation. The role of ORF I or of a ORF I gene product in tyrosinase expression remains however unknown.

The leader sequence of the ORF I protein was able to direct expression and subsequent secretion of the tyrosinase from the filamentous fungus *Neurospora crassa* in *S.*

glaucescens and *S. lividans* host strains, be it under the control of the *mel* or the *aph* promoter. The 33 amino-terminal amino acid residues of the ORF I protein provided sufficient information to direct export of this cytoplasmic protein across the cytoplasmic membrane and cell wall of *Streptomyces*. A shorter version of the leader peptide containing only 27 amino acid residues failed to direct secretion of this protein. Cleavage of the leader peptide between the amino acid residues Ala(32)-Ala(33) appears to be most likely.

IA. ZUSAMMENFASSUNG

Das Melaninoperon von *Streptomyces glaucescens* besteht aus drei offenen Leserastern, welche als eine polycistronische mRNA transkribiert werden. Das zweite offene Leseraster entspricht dem Strukturgen für das Enzym Tyrosinase (*melC*). Ein 402 bp grosses offenes Leseraster (ORF I), das eine mutmassliche Signalpeptidsequenz enthält, liegt unmittelbar vor dem *melC* Gen. ORF I kodiert für ein Protein, welches sowohl im Kupfertransport als auch bei der Sekretion von Tyrosinase beteiligt sein könnte.

Die Nukleotidsequenz eines 1489 bp langen DNA-Fragmentes, welches unterhalb des *melC* Gens liegt und das dritte Leseraster (ORF III) des *mel*-Operons enthält, wurde bestimmt. Der ORF III weist den für *Streptomyces* typischen einseitigen Gebrauch von Aminosäurecodons auf, der Codons mit G oder C an der dritten Position stark bevorzugt, und kann für ein 21'797 Dalton Protein kodieren. Es konnte jedoch kein Hinweis für die Existenz eines Proteins gefunden werden, welches durch ORF III kodiert wird.

Die Startstelle der Transkription des *mel*-Operons wurde mit S1-Nuklease Kartierung sowie mit "Primer"-Verlängerung bestimmt. Sie liegt 32 bis 34 bp oberhalb des Translations-Startcodons des ersten offenen Leserasters (ORF I) des Operons.

Die Transkriptions-Terminationsstelle des Operons wurde mit S1-Nuklease Kartierung bestimmt. Sie befindet sich 2056 oder 2057 bp nach dem Transkriptionsstart (+1) des Operons. Der Transkriptionsendpunkt liegt am Ende des Stammes einer Haarnadelstruktur in der DNS, die eine freie Bildungsenergie (ΔG^0) von -29kcal aufweist. Es wird deshalb ein Terminationsmechanismus vorgeschlagen, welcher dem der Rho-unabhängigen Termination in *E. coli* entspricht.

Zwischen 172 und 219 bp der DNS Sequenz oberhalb des Transkriptionsstartpunktes sind notwendig für einen vollständig aktiven und regulierten *mel* Promotor. Deletionsanalyse, Gel "retardation"- und DNase I- "footprint" Experimente ermöglichten die funktionelle Aufteilung der Promotorregion in eine RNA Polymerase Erkennungsstelle bis zur Position zwischen den Nukleotiden -33 und -42, eine Bindungsstelle für ein Protein mit wahrscheinlich regulatorischer Funktion zwischen Position -65 und -93 sowie eine "upstream" Aktivierungsstelle (UAS), welche zwischen den Nukleotidpositionen -158 und -219 liegt.

Der oberhalb des *melC* Gens gelegene ORF I beeinflusst die Expression des Tyrosinasegens. Deletionen und "nonsense"-Mutationen in der Sequenz des ORF I erzeugten eine starke Abnahme oder den totalen Verlust der Tyrosinaseaktivität. Der Effekt der Deletionsmutationen konnte durch eine intakte Kopie des ORF I in *trans*

komplementiert werden, während die "nonsense"-Mutationen nicht komplementiert werden konnten. Es bleibt jedoch unklar, welche Rolle der ORF I oder das ORF I-Produkt bei der Expression der Tyrosinase spielt.

Die Leadersequenz des ORF I Proteins ermöglichte Expression und anschließende Ausscheidung der Tyrosinase aus dem filamentösen Pilz *Neurospora crassa* in den beiden Wirtsorganismen *S. glaucescens* und *S. lividans*, sowohl unter der Kontrolle des *mel*-Promotors als auch des *aph*-Promotors. Die 33 aminoterminalen Aminosäuren des ORF I Proteins enthielten die ganze Information für die Ausscheidung dieses zytoplasmatischen Proteins durch die Zytoplasmamembran und die Zellwand von *Streptomyces*. Eine kürzere Version des Leaderpeptides, die nur 27 Aminosäuren enthält, konnte hingegen dieses Protein nicht sekretieren. Die Signalpeptidase schneidet das Signalpeptid wahrscheinlich zwischen den Aminosäuren Ala(32)-Ala(33).