Adam Hopkins, Grant Buchanan and Tracy Palmer.

The role of the twin arginine protein transport pathway in the assembly of the *Streptomyces coelicolor* cytochrome bc_1 complex

Supplementary Information.

*Xho*I *Bam*HI
ctcgaggatccaccgccttccgccggaacggcggggtccgggcacgccaaacccctcctgtggctgtggccggcc

accgccgtcaccttcggaccccgtggagccgctcccggttccacggggtccgaaggtgtgatgagcaggctgcgc

cttcctcgcgcggccgcaaggtacgagttgatgaccttgtttatccgcatctgaccaattttgatcgcttacggg

gtgtgactcgggccacgcggattgggcgtaacgctcttgggaacaacacgatgacctaagaggtgacagccgcgg

agggaatacggacgccgttcacggcgctgtgcatctccccggcccgcccgcaccgtcggcccattcccaagccgg

tggtcggcccctgtccgccgtggacggggccggaagccgtttttcaacgttccgagaggttgttccat**ATG**aaaa

agctctccgcacgacgacgccatccgctggcggccctcgtcgtcctactcctcgcgctggcatgcacagggggc

tgtacgccgcgttcgcacccgcgagcaaggcgcaggccgatgaatccgcccagtccctcgccatcgacgagggca

agaagctctacgccgtcggctgtgccagttgccacggcaccggagggcagggcaccagtgacggtccgagcctgg

tcggcgtaggcgccgcggccgtcgacttccaggtgggcaccggccgcatgccggcccagcagcccggcgcgcagg 750

tcccgaagaagaaggtcatctactcccaggcggagatcgaccagctcgcggcgtacatcgcgtcgctgggcgccg

gtccggcgatcccctcggaggagaagtacggccccgagggcgcggacatcgccaagggtggtgagctgttccgca

ccaactgcgcgcagtgccacaacttcaccggcaagggcggcgccctgacgcacggcaagtacgcgccgagcctGg

agggtgtcgacccgaagcacatctacgaggccatgcagaccggcccgcagaacatgccctccttccccgacacca

cgctgtcggagcagaacaagaaggacatcatcgcctacctggacgcggtcaacggtgacgacaccgagagccccg

gtggtctcagcctcggcggactcggcccggtcagtgagggtctgttcgcctgggtgttcggactgggcgcgctga

*Bsi*WI
tcgccgtcgccgtctgggtcgccgctcgTacGgcaaaggccaagaagtcaggc*caccaccaccaccaccaccacc*

*Bsi*WI
*accaccac***tga**gtcgccgctcgTacGgcaaaggccaagaagtc**ATGa**gtagccaagacattccagaagagaacct

gccggcagagcaggaccgcccgcacggcgcggccgcccggcccgcggacgagaccaacccgttcgccgacccggg

tctgccgccccacgagccgcgggtccaggacgtcgacgagcgggccgccaagcggtccgagcgcacggtcgccct 1500

gctgttcacgctgtcgatgctggccaccatcgccttcatcgccgccttcgtggcgatcgacgtcgacaagtcggt

ctacatcttcccgctcggtcacatcagtgcgctgaacttcgcgctcggcatgacgctcggcgtcgcgctgttcgc

catcggcgcgggcgcggtccactgggcccgcaccctgatgtccgacgaggaggtcgccgacgagcgtcacccgat

cgaggcgtcccccgaggtccgtgccaaggtccacgcggacttcaagcagggtgccaaggagtccgtgatcgggcg

ccgcaagctgatccgcaacacgatgctgggcgcgctcaccctggtgccgctctccggcgtcgtcctgctgcgcga

cctcggtccgctgcccgggaccaagctccgccacaccctgtggtccaagggcaagctcctcgtcaacatgaacac

caacgagccgctgcgtccctccgacgtcgcggtgggctcgctcaccttcgccatgcccgagggcctggaggagca

cgacgaggacttccagaacgagatcgccaaggccgccctgatgatcatccggctggagccggactccatcaagga

caagcgcgagctGgagtggtcgcacgagggcatcgtggcgtactcgaaAatctgcacccacgtcggttgcccgat

ctccctgtacgagcagcagacgcaccacgcgctctgcccctgccaccagtccaccttcgaccttgccgacggtgc 2250 ccgggtcatcttcggtcccgccggtcacgccctgccgcagctgcgcatcggcgtgaacgacgagggttacctcga

*Age*I
agcgctcggcgacttcgaggaAccGgtcggtcctgcttactgggagcgcggaggc*caccaccaccaccaccacca*

*Age*I
*ccaccaccac***tga**gaggaAccGgtcggtcctgcttactgggagcgcgg**ATGa**gtactgcagcaaacgaaccgtcc

cgctcgcgcgggaaggcaccggccggcgagcgcgtcgccgactgggccgacggccggctggggatctactccctg

gccaaggccaacatgcgcaaAatcttccccgaccactggtcgttcatgctgggtgaggtctgcctctacagcttc

atcatcatcatcctcacgggtgtgtacctgacgctgttcttccacccgtcgatggccgaggtcgagtaccacggc

tcCtacgtgccgctgcagggccagatgatgagtgaggcctacgcctccactctggacatcagcttcgacgtccgc

ggcggtctgctgatccggcagatccaccactgggccgcgctgatcttcctggccggcatgttcgtgcacatgatg

cgcgtcttcttcaccggcgcgttccgcaagccgcgcgaggtcaactggctgttcggcttcctgctgctcgtcctc

ggcatgttcaccggcttcaccggAtactcgctcccggacgacctgctctcgggcaccggAatccgcttcatggag 3000

ggcgcgatcctgtccgtgccgatcgtcggcacgtacatctcgttcttcctgttcggcggcgagttccccggccac

gacttcgtgtcccggttctactcgatccacatcctgctgctgccgggcatcatgctcgggctgctggtgggccac

ctgatcctggtcttctaccacaagcacacgcagttcgcgggtcccggcaagaccaacaagaacgtcgtcggcatg

ccgctgctgccggtgtacacggcgaaggccggaggcttcttcttcctggtcttcggtgtgatctccgtcgtgtcg

gccatcgccacgatcaacccgatctgggccatcgggccctaccggcccgaccaggtctccaccggcgcccagccg

gactggtacatgggcttctccgagggcctgatccgggtgatgcccggctgggagatcaacgcctggggtcacaca

ctcgtcctgggcgtgttcgtcccgctgctgatcttcccgctggtcctggcggcgatcgcggtctacccgttcatc

gagtcctgggtcaccggcgacaagcgcgagcaccacatcctggaccgcccgcgcaacgccccgacccgtacCgcc

ttcggtgtcgcctggctgaccgtctacttcgtgctgctgatcggtggcggcaacgacctgtgggccacccacttc

cacctgtcgatcaacgcgatcacctggttcgtccgcatcgcgttcttcgtcggTccggtcgtcgccttcatcgcc 3750

accaagcggatctgcctcggcctccagcgccgggacaaggacaaggtgctgcacggacgcgagtccggcatcatc

aagcgcctgccgcacggtgagttcatcgaggtgcacgagccgctcagccaggagcagctgcacacgctcacggcg

cacgagcagtaccagccggcggagatcggcccgacggtcgacgagaacggcgtcgagcgcaaggtgagcggcacg

cagaagctccgtgccaagctcagcgagtcgtactacggcgaggagtcgcagattccgaagccgaccgtcgaggag

*Nco*I *Eco*RI
tacaaggagatcacgagcggccaTggccaccacggc*caccaccaccaccaccaccaccaccac***tga**gaattc

*Xba*I *Nco*I *Eco*RI *Xba*I *Sac*I
cccttctagaccaTggccaccac**tga**gaattccccttctagagctc 4171

Figure S1. Sequence of the original synthetic construct encoding *qcrCAB*. Restriction enyzme sequences are in underline with the name of the restriction enzyme given above. Start codons for each gene are capitalized and shown in bold underline, stop codons are given in bold lower case. Deca-histidine coding sequences are shown in italic underline. Silent mutations in the sequence to remove or add restriction sites are capitalized plain font. The *hrdB* promoter region covers nucleotides 12 – 443, the $qcrC_{his10}$ coding sequence covers nucleotides 444 – 1286, the $qcrA_{his10}$ coding sequence covers nucleotides 1319 – 2416 and the $qcrB_{his10}$ coding sequence covers nucleotides 2449 – 4119.

QcrA

1 **mssqdipeen lpaeqdrphg aaarpadetn pfadpglpph eprvqdvder a**akrsertva

61 llftlsmlat iafiaafvai dvdksvyifp lghisalnfa lgmtlgvalf aigagavhwa

121 **rtlmsdeeva derhpieasp evrakvhadf kq**gakesvig rrkli**rntml galtlvplsg**

181 **vvllrdlgpl pgtklrhtlw skgkllvnmn tneplrpsdv avgsltfamp egleehdedf**

241 **qneiakaalm iirlepdsik dkrelewshe givayski**ct hvgcpislye qqthhalcpc

301 hqstfdladg a**rvifgpagh alpqlrigvn degylealgd feepvgpayw erg**

QcrB$_{His10}$

1 **mstaanepsr srgkapager vadwadgrlg iyslakanmr k**ifpdhwsfm lgevclysfi

61 iiiltgvylt lffhpsmaev eyhgsyvplq gqmmseayas tldisfdvrg gllirqihhw

121 aaliflagmf vhmm**rvfftg afrk**prevnw lfgflllvlg mftgftgysl pddllsgtgi

181 rfmegailsv pivgtyisff lfggefpghd fvsrfysihi lllpgimlgl lvghlilvfy

241 h**khtqfagpg kt**n**knvvgmp llpvytaka**g gffflvfgvi svvsaiatin piwaigpyrp

301 dqvstgaqpd wymgfsegli rvmpgweina wghtlvlgvf vpllifplvl aaiavypfie

361 swvtgdk**reh hildrprn**ap trtafgvawl tvyfvlligg gndlwathfh lsinaitwfv

421 **riaffvgpvv afiatkricl glqrr**dkdkv lhg**resgiik rlphgefiev heplsqeqlh**

481 **tltaheqyqp aeigptvden gverkvsgtq klraklsesy ygeesqipkp tveeyke**its

541 ghghhghhhh hhhhhh

QcrC

1 mkklsarrrh plaalvvlll alactgglya afapaskaqa desaqslaid egkklyavgc

61 aschgtggqg tsdgpslvgv gaaavdfqvg tgrmpaqqpg aqvpkkkviy sqaeidqlaa

121 yiaslgagpa ipsee**kygpe gadiakg**gel frtncaqchn ftgkggalth g**kyapslegv**

181 **dpkh**iyeamq tgpqnmpsfp dttlseqnkk diiayldavn gddtespggl slgglgpvse

241 glfawvfglg aliavavwva artakakks

Figure S2. Identification of purified cytochrome *bc*$_1$ complex components by tryptic peptide mass fingerprinting. Recovered peptides are given in red, while predicted transmembrane helices are underlined. The coverage QcrA, for His-tagged QcrB, the Rieske protein and QcrC are 36%, 67% and 9%, respectively.