UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Identification of additional homologues of subunits VII and VIII of the ubiquinol-cytochrome c oxidoreductase enables definition of consensus sequences

Boumans, H.; Berden, J.A.; Grivell, L.A.

Link to publication

Identification of additional homologues of subunits VII and VIII of the ubiquinol–cytochrome c oxidoreductase enables definition of consensus sequences

Hans Boumans[a], Jan A. Berden[a,*], Leslie A. Grivell[b]

[a]E.C. Slater Institute, University of Amsterdam, Plantage Muidergracht 12, Amsterdam 1018 TV, The Netherlands
[b]Department of Molecular Cell Biology, Section for Molecular Biology, University of Amsterdam, Amsterdam, The Netherlands

Abstract The *Candida utilis QCR7* gene encoding subunit VII of the ubiquinol–cytochrome c oxidoreductase was isolated by functional complementation of the *Saccharomyces cerevisiae* subunit VII-null mutant. Several other subunit VII homologues as well as homologues for subunit VIII were identified by screening the GenBank database. Some of these homologues for subunit VII could only be identified as such using a consensus sequence that was derived from the multiple sequence alignment. Definition of the consensus should facilitate further analysis of structure/function relationships in this protein.

Key words: Ubiquinol–cytochrome c oxidoreductase; Multiple sequence alignment; Consensus sequence; *Candida utilis*

1. Introduction

The ubiquinol–cytochrome c oxidoreductase (bc_1 complex) plays a central role in the respiratory chain, transferring electrons from ubiquinol to cytochrome c [1]. Coupled to this reaction is the translocation of protons across the inner membrane of mitochondria in eukaryotic organisms and across the cytoplasmic membrane in many bacteria. The bc_1 complex of these prokaryotes consists of 3 or 4 subunits [1] while those from eukaryotes may consist of up to 11 subunits [2]. Three subunits, cytochrome b, cytochrome c_1 and the Rieske Fe-S protein, contain prosthetic groups and are present in all known bc_1 complexes. Since both prokaryotic and eukaryotic bc_1 complexes function in a basically similar fashion, the function of the so-called supernumerary subunits is a matter for speculation.

For the bc_1 complex from *Saccharomyces cerevisiae* 10 subunits have been described until now [3,4]. Both subunit VII and VIII (14 kDa and 11 kDa, respectively) are thought to be in close association with cytochrome b [5], the core subunit of the complex, and a role in the binding of ubiquinol/ubiquinone has been suggested [6–8]. Disruption of the genes encoding these subunits showed that both are essential for correct assembly of the complex [9,10].

Here we report on the isolation of the subunit VII homologue of *Candida utilis* by functional complementation and the identification of a number of other homologues of both subunits VII and VIII by database searches (Table 1). Multiple sequence alignment permits definition of a consensus sequence for both

proteins which should prove to be of value in the identification of new members of the two families and in the design of mutant subunits for study of structure/function relationships.

2. Materials and methods

Escherichia coli strain JF 1754 (*lac, gal, metB, leuB, hisB, hsdR*) was used for DNA manipulations throughout. *E. coli* transformants were grown in YT medium (1.6% (w/v) bactotryptone, 1% (w/v) yeast extract and 0.5% NaCl) containing 200 μg/ml ampicillin. *S. cerevisiae* strain DLL70 (*a, his3, ura3, leu 2; qcr7::LEU2*) [9] was used for the transformation of plasmids. Transformation of yeast was performed according to [11]. Transformants were selected on minimal media containing 0.67% (w/v) yeast nitrogen base (Difco), 2% (w/v) glucose, 2% (w/v) agar supplemented with histidine (20 mg/ml). The respiratory capacity of the transformants was checked on solid media containing 1% (w/v) yeast extract, 1% (w/v) bactopeptone, 2% (v/v) glycerol and 2% (v/v) ethanol. Isolation of plasmid DNA from yeast was performed according to [12].

Genomic DNA was isolated from *C. utilis* CBS 621 and partially digested with *Sau*IIIA. Fragments were separated by electrophoresis. Fragments ranging from 3.5 to 15 kb were isolated from agarose gel and ligated into the *Bam*HI site of YepLac195 [13]. Transformation into *E. coli* resulted in about 75,000 transformants. Over 95% of these contained the YepLac195 vector with an insert. Transformants were pooled and plasmid DNA was isolated.

An GCG sequence analysis package was made available by the CAOS/CAMM center in Nijmegen, The Netherlands.

3. Results

3.1. Isolation and sequencing of C. utilis QCR7

The *C. utilis* genomic library was transformed into the *S. cerevisiae QCR7*-null mutant DLL70 [9]. Of the approximately 10,000 transformants that were replica-plated on ethanol/glycerol (EG) medium, four complemented the respiratory deficiency of the DLL70 strain. Plasmid DNA was isolated and amplified in *E. coli*. Restriction analysis showed that three clones were identical while the fourth, larger clone contained an overlapping DNA insert. The size of the smallest clone was approximately 7000 bp and a restriction map for enzymes present in the multiple cloning site of YepLac195 is shown in Fig. 1. Three subclones were constructed using the *Sst*I restriction sites and subsequently transformed into DLL70. Subclone 1, containing the first 1750 bp, was able to suppress the EG-negative phenotype of DLL70.

Suitable fragments were subcloned into vector pBluescript SK (Stratagene) and the DNA sequence was determined on one strand only [14]. Only one open reading frame (ORF) longer than 200 bp was found. The DNA sequence of the ORF plus some 400 bp of flanking sequence was determined on the other

*Corresponding author. Fax: (31) (20) 525-5124.

Table 1
List of sequences and their accession numbers which are identified in this study

Organism	Homologous to	Accession no.
C. utilis	subunit VII	U20215
D. melanogaster		X67704
O. sativa (rice)		D23346/D23977/D24061
P. falciparum		T18013/T18091
C. elegans		T01178
Z. mays (maize)		T27569
A. thaliana		T42220/T42225
O. sativa (rice)	subunit VIII	D15978/D38788/D39007
A. thaliana		T22556

Sequences can be found in GenBank, EMBL and the DNA Data Bank of Japan.

Fig. 1. Restriction map of the DNA-region containing the QCR7 gene of C. utilis. Direction of transcription is indicated by arrow. Only relevant restriction sites are shown. Restriction sites below the line are derived from the polylinker. Subclones made using the SstI restriction sites are indicated.

strand as well. This sequence together with the translated ORF is depicted in Fig. 2.

The QCR7 gene of C. utilis is 378 bp long and codes for a protein of 126 amino acids with a calculated molecular weight of 14.1 kDa. Alignment with the S. cerevisiae subunit VII shows 54% sequence identity, which is more or less equally distributed over the entire protein. Secondary structure predictions and hydrophilicity plots also indicate a high level of conservation between the two proteins (data not shown).

The capacity of DLL70 complemented with the C. utilis QCR7 gene to grow on a non-fermentable carbon source indicates that the heterologous bc_1 complex in which subunit VII is replaced by its homologue from C. utilis is (at least partially) functional. However, whether lower efficiency is due to sequence differences or lower expression of the C. utilis gene in S. cerevisiae still has to be resolved.

3.2. Database search for subunit VII homologues

We have performed a database search for proteins or ORFs showing sequence identity with subunits VII and VIII. A TBLASTn search was performed using yeast subunit VII as the query sequence. Of the sequences retrieved only those with a Smallest Sum Probability $P(N)$ smaller than 10^{-3} were considered to be significantly similar. In addition to the known homologues of this protein a Drosophila melanogaster clone was found with sequence similarity of statistical significance $(P(N) = 8.2\ e^{-6})$. This clone contains the Mst98Ca and Mst98Cb genes. Downstream of these genes is an ORF encoding part of subunit VII. According to the multiple sequence alignment in Fig. 3 about 75 bp at the 3'-end of the gene are missing.

In another TBLASTn search we used the Solanum tuberosum (potato) subunit VII as the query sequence. This resulted in the retrieval of 3 clones containing rice (Oryza sativa) cDNA with significant similarity. These clones were expressed sequence tags (EST) and each contained only part of the QCR7 gene. However, when aligned these 3 sequences permit reconstruction of a complete QCR7 gene. The overlapping regions showed complete identity, suggesting that these cDNAs were translated from the same gene. In addition, one maize (Zea mays) cDNA was retrieved containing part of the QCR7 gene. Several ambiguous nucleotides were present and an additional ambiguity had to be introduced at position 99 in order to align the translated sequence with the potato protein. Two clones containing Arabi-
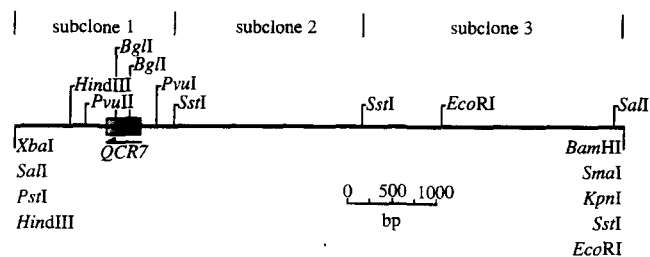
dopsis thaliana cDNA were also found. Both sequences contained several ambiguous nucleotides, but no differences at the DNA level were found. Combination of the two sequences permitted elimination of all but one of the ambiguities.

The homologues of subunit VII known at this stage (from S. cerevisiae [15], Kluyveromyces lactis [16], C. utilis (this study), human [17], rat [17], bovine heart [18], D. melanogaster (this study), S. tuberosum [19], Triticum aestivum (wheat) [20], rice (this study), A. thaliana (this study) and Z. mays (this study)) were aligned using the GCG PileUp program [21]. This alignment permitted the generation of a consensus sequence containing 15 amino acids that are completely conserved. Using this consensus sequence we re-examined the sequences found in the two previous TBLASTn searches which had a statistical significance of $P(N) > 10^{-3}$. In this way we could identify part of subunit VII of Caenorhabditis elegans and part of the homologue from Plasmodium falciparum. Alignment of these two protein sequences with S. cerevisiae subunit VII using the MACAW program [22] confirmed this homology. Between the S. cerevisiae and P. falciparum proteins a block was found of 39 amino acids (residues 39–77 in the S. cerevisiae protein) with $P = 4.6\ e^{-7}$. Between S. cerevisiae and C. elegans subunit VII a block of 39 amino acids was found (residues 77–115 in the

```
-297                    CCCTCCCACCATCGGAATTGTTCTGGC
-270 GATGCTGTCGGAATTAGCATCCGTCATACACAATTTCGGAACACC
-225 CCGAAAACGGCAGAACTTCCACAACAAGCGATCGCGCGGCCGGGT
-180 AACCGCAGAGTCGCCGCGGCCAATGAAACAGCGGGAAAAACAAAT
-135 CGCGAAGCTCTCTTAGTGCGGGAGCCCCCGATGTACCAAAGCAAG
 -90 GGCAGAGAAAAAAAACCACGATAAAAAAAGCGAACTGGCCTAATA
 -45 GAAGAGCACAGTTTTACATCAACAAACCCATTCAAGACAAAAACA
   1 ATGGCTTCCCACACCTCCATCGTCAAGGCCGGTGACTTCATCCTG
     M  A  S  H  T  S  I  V  K  A  G  D  F  I  L
  46 CGTACCCCAGTGCTTAGATCCATCTTCGTGCCAGCCGCCAAGTTG
     R  T  P  V  L  R  S  I  F  V  P  A  A  K  L
  91 TTCTGCGCTTACTCTGGCTACCGCCAGCTCGGCCTCAAGTTTGAC
     F  C  A  Y  S  G  Y  R  Q  L  G  L  K  F  D
 136 GACTTGATCCACGAGGAGAACCCAACCGTTCAGAAGGCGCTTTCC
     D  L  I  H  E  E  N  P  T  V  Q  K  A  L  S
 181 AGATTGCCAAAGGACGAGATCTACGCCAGAAACTTCAGAATGCTG
     R  L  P  K  D  E  I  Y  A  R  N  F  R  M  L
 226 ACTGCTGCTCAGTGCGGTATCACCCACCACTTGCTTTCCGCCGAC
     T  A  A  Q  C  G  I  T  H  H  L  L  S  A  D
 271 AAGGCTCTTAAGCCATCCGAGGACACTCCGTACTTGTTGCCATAC
     K  A  L  K  P  S  E  D  T  P  Y  L  L  P  Y
 316 TTGTTGGAACTTGAGGCCGAGGCTGCTGAGAGAGTTGAGCTGGAT
     L  L  E  L  E  A  E  A  A  E  R  V  E  L  D
 361 AACGTTGAGGTTGCCAAATAAGCTCTTTCATGAGGAGCAATGGTA
     N  V  E  V  A  K  *
 406 CGTTGAGGCTACAAGGAGCGAGGAGCAGCCTCTTTCACATTCACA
 451 GTACAAGTTATATGAGTCTAAAATAAACAAAAAAA
```

Fig. 2. DNA sequence and deduced amino acid sequence of the QCR7 gene and flanking sequences. The stop codon is indicated *.

```
1   MPQSFTSIARIGDYILKSPVLSKLCVPVANQFINLAG...YKKL
2   MPQTFTSIAKIGDYILRTPALAKVVVPIAHQFINLSG...YRKM
3   MASHTSIVKAGDFILRTPVLRSIFVPAAKLFCAYSG...YRQL
4          MAGKQAVSAS..GKWLDGIRKWYYNAAG...FNKL
5               R..SKWLDGFRKWYYNAAG...FNKL
6          AGRPAVSAS..SRWLEGIRKWYYNAAG...FNKL
7          MSKYVARVG..PAVFSKLGKWAYNMSG...FNQY
8             HKLIFPYYFKFIRAPYERWQFCATTKFLREH
9       MLSAFSQWFVNPRRNPLARXHMQAISSX...LRKY
10      MASSFSRWLVDPKKNPLAAIHMKTLSSR...LRNY
11                                         KNY
12
13
14        MASLLKAFIDPKKNFLARMHMKAISTR...LRRY
```

```
1   KF LIAEENPI.MQT LRR EDESYA AY IIRAHQTELTHHL
2   RF LIEEENEL.AQT LRR ADESYA IY IINAHQLSLSHHL
3   KF LIHEENPT.VQK LSR KDEIYA NF MLTAAQCGITHHL
4   MR TIYEDEDV.KE. IRR ENLYND MF IKRALDLNLKHQI
5   MR TMHETEDV.KE. IRR ENLYND MF IKRALDLSMRHQI
6   MR TIHENDDV.KE. IRR ENLYDD VF IKRALDLSMRQQI
7   YR CLYENEDV.AE. VRR RKLYDE NY ILRALHLSMTKTI
8   MY MYSDKDPV.IER ISL KDIQTR YR MLRGTHINYLRLF
9   RY LYDPKHDLDIKE LER REVVDA HQ LKRAMDLSMKHQY
10  RH LYDPMYDLDVKE LNR REIVDA NQ LLRAMDLSMKHQY
11  RY LYDPYFDLDIKE LGR REVV          KHQY
12                         SRAHTLALHGEK
13               LGR REVVDX IQ LKXAMDLSMKHQY
14  RY LYDQYYSMDIKE MNR REVVDA NQ LKRAMDLSMKHEY
```

```
1   PRNQWIKAQEDVP LPYILEAEAAAKEKDELDNIEVSK
2   PKDKWTKPEDDIP TPYLLEAEAFVKEKEELDNLEVAK
3   SADKALKPSEDTP LPYLLELEAEAAERVELDNVEVAK
4   PKEQWTKYEEENF EPYLKEVIRERKEREEWAKK
5   PKDQWTKYEEDKF EPYLKEVIRERKEREEW
6   PKEQWTKYEEDKS EPYLKEVIRERKEREEWAKK
7   PKEQWTKYEEDIK
8   HPSE.QNYDPYIP
9   SENDQAQQTPFRG SDMMDLVKKERLEREELGALPLHQRTLP
10  EEDLQAMQTPFRN QEMLALVKRESAEREALGALPLYQRTLP
11  PDDVRALQQPFKS SDMLALVKK    ALGALPLYERTLP
12  PKAEWTQWDQESW KPYLDEXEAEKKLAPRHXDSFHHTK
13  PXXLQAQQVPFRG GDMLALVKKE
14  PKXLQAVQTPFRG QDMLALVERGSKEREALGSSYHLLPADNSH
```
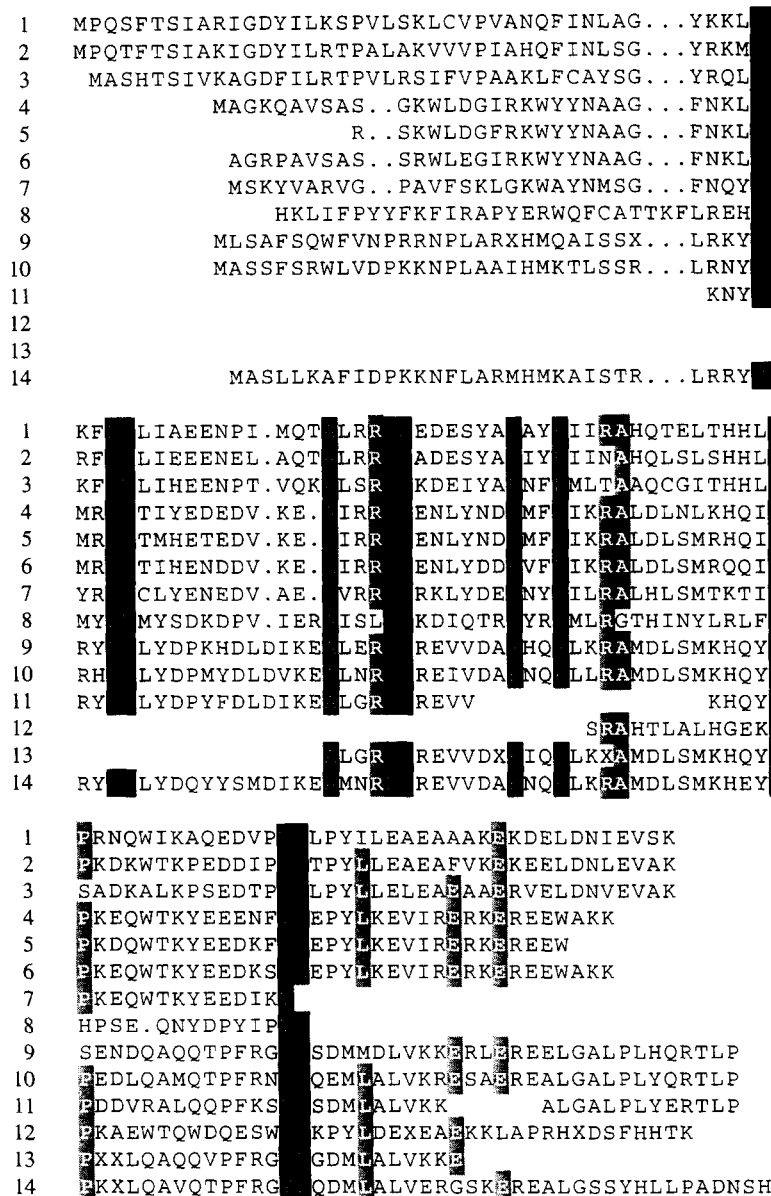
Fig. 3. Mulitple sequence alignment of the subunit VII homologues. Sequences were aligned using the GCG PileUp program (gap weight, 3.00; gap length weight, 0.10) [21]. Amino acids conserved in all sequences are indicated by a black box; amino acids appearing in at least 75% of the sequences are indicated by a grey box. Gaps are indicated (.). 1 = S. cerevisiae [15]; 2 = K. lactis [16]; 3 = C. utilis (this study); 4 = human [17]; 5 = rat [17]; 6 = bovine heart [18]; 7 = D. melanogaster (this study); 8 = P. falciparum (this study); 9 = O. sativa (this study); 10 = S. tuberosum [19]; 11 = T. aestivum [20]; 12 = C. elegans (this study); 13 = Z. mays (this study); 14 = A. thaliana (this study). Note that sequences from 5, 7, 8, 11, 12 and 13 are only partial.

S. cerevisiae protein) with $P = 8.7 \, e^{-7}$. Two clones were found for P. falciparum which were almost identical. Since several ambiguous nucleotides are included by the authors it is likely that the differences between the two clones are due to sequence errors. The homologues from these two organisms clearly match the consensus sequence (see Fig. 3). However, the statistical value $P(N)$ given by TBLASTn was 0.40 for P. falciparum and 0.72 for C. elegans.

The multiple sequence alignment of all subunit VII homologues shows a clear consensus sequence. The SwissProt protein database was screened for proteins containing a similar pattern using a profile analysis approach [23]. The multiple sequence alignment in Fig. 3 was used as an input to make a

profile by the program PROFILEMAKE. This profile is than used by the program PROFILESEARCH which screens the database [23]. Unfortunately, no protein was found having significant similarity to the profile of the subunit VII family.

3.3. Database search for subunit VIII homologues

Similar search strategies were used to identify new members of the subunit VIII family. Using the potato subunit VIII as a query sequence resulted in the identification of the complete QCR8 gene of A. thaliana. The sequence contains several ambiguous nucleotides and is 13 amino acids longer than the potato protein. The proteins differ extensively at the C-terminus, in clear contrast to the rest of the protein. However, dele-

```
1     MGPPSGKTYMGW.WGHMGGPKQ GITS AV  YA KPLQGIFHNAVFNS
2    MGGPHAKAYMGW.WGSIGSPAQ GITT TV  YA KPLNNIFHNAVFNT
3  MGGAAGGKTYLGW.WGHLGGPKQ GIIT SL  FQ RPMAGFFKTSTQNM
4              GRQFGHLTRVRH..VIT SL  FE RAFPHYFSKGIPNV
5                  MGKTPVRM AVV. AL  FQ KVMPGLWKDITTKI
6                  MGKQPVKL AVV. AI  FQ KIMPGLWKDLPGKI
7                  MGKQPVKX AVV. ALP FQ XIXTGLWKXLPEKI


1  FRRFK....SQF YVLIP.AGIYWYWWKNGN YNEFLYSKAGREELERVNV
2  FRRVK....SQI YMALP.AALYWAWWVNCRDYNAYLYTKAGREELERVNV
3  FRRVM....TEG YVAIP.FGIAYYIYCWGK RNEFLNSKHGRHLVEE
4  LRRTR....ACI RVAPPFVAFYLVYTCGTQ FEKSKRKNPAAYENDR
5  HHKVSENWIFATVLLGPIVGTYEYAMYYKEQ KLFHRY
6  HHKVSENWISAT LLGPLVGTYSYVQHFLEK KLEHRY
7  HHKVSXNWXSAT LVTPVVGTYWYAQYFKEQ KLLHRF
```

Fig. 4. Mulitple sequence alignment of the subunit VIII homologues. Sequences were aligned using the GCG PileUp program (gap weight, 3.00; gap length weight, 0.10) [21]. Amino acids conserved in all sequences are indicated by a black box; amino acids appearing in at least 75% of the sequences are indicated by a grey box. Gaps are indicated (.). 1 = *S. cerevisiae* [31]; 2 = *K. lactis* [32]; 3 = *S. pombe* [27]; 4 = bovine heart [33]; 5 = *O. sativa* (this study); 6 = *S. tuberosum* [34]; 7 = *A. thaliana* (this study).

tion of T_{225} and the consequent introduction of a frame shift in the coding region, results in a clear improvement of sequence match at the C-terminus and reduces the length of the protein to that of the potato homologue. As a large number of ambiguous nucleotides was introduced in this sequence, a sequence error is not surprising.

In addition three clones were retrieved from *Oryza sativa* (rice). These clones are expressed sequence tags and all contain the complete *QCR8* gene. Similar to *A. thaliana*, two of these clones contained two ambiguous nucleotides and a frame shift was found in one of them (A_{244} had to be deleted in clone C0237-1A). Alignment of the three genes showed that two of them were identical while the third sequence showed six differences on the nucleotide level resulting in four amino acid changes. This third sequence is displayed in the sequence alignment in Fig. 4 as it did not contain any ambiguous nucleotides.

Although the multiple sequence alignment of the subunit VIII homologues lacks a clear consensus sequence, a profile was generated and used to search the SwissProt database similar to the approach for subunit VII. Also in this search no proteins were found with significant similarity.

4. Discussion

We have isolated a gene from *Candida utilis* and identified it as the *QCR7* gene encoding subunit VII of the bc_1 complex based on sequence identity with the *S. cerevisiae* subunit VII and on the fact that a plasmid carrying this gene is able to functionally complement the *S. cerevisiae* *QCR7*-null mutant.

The second part of this study involves the identification of new homologues of both subunits VII and VIII. These homologues were identified by screening DNA and EST databases using protein sequences of several known subunit VII and VIII homologues. Not all of the sequences found in this study could be identified based on sequence identity with known proteins. Here we show that the use of a consensus sequence as a tool

in screening sequence databases increases the chance of identifying homologous proteins.

The conserved amino acids, indicated by a black box in Fig. 3, are not distributed over the entire protein, but are predominantly found in the central domain (residues 31–89 in *S. cerevisiae*). The N-terminus (residues 1–30 in *S. cerevisiae*) is highly variable, both in amino acid composition and in length. The N-terminus of *S. cerevisiae* subunit VII (residues 1–18) has been proposed to contain a mitochondrial targeting sequence [24] as it may form an amphipathic helix with a high hydrophobic moment. Apart from the N-terminal methionine, this targeting sequence is not cleaved [1].

The C-terminus (residues 90–127 in *S. cerevisiae*) contains only a few conserved amino acids. However, the sequences do show additional similarity in that they all contain a large number of both acidic and basic residues, making this part of the protein highly hydrophilic. Another interesting feature of this region is that five positions in the alignment contain more than 50% glutamate residues. A previous study on this C-terminus in *S. cerevisiae* showed the importance of this part of the protein for correct assembly of the bc_1 complex, but not for electron transport [25]. The study also implied that the overall features of this region (hydrophilicity and charges) rather than specific residues are important, a feature which is confirmed by the sequence alignment in Fig. 3.

Three clones were found containing rice *QCR8*. Two were identical while the third contained changes both in nucleotide and amino acid sequence. Although sequence errors could be responsible for these differences, it should be noted that gene duplicates are relatively common in plants. In the case of the Rieske Fe-S protein it was shown that in tabacco at least 5 different genes are present and transcribed [26]. Most of these proteins differ only by a few amino acids. For subunit VII also three clones from rice were identified. In this case no differences were found between the different clones. These clones were, however, only partly overlapping.

The multiple sequence alignment in Fig. 3 shows that there is a clear consensus sequence for subunit VII. This subunit VII 'family' now contains members of the classes of yeasts, mammals, plants, insects, protozoa and eumetazoa. In marked contrast to this, is the low sequence identity amongst the subunit VIII homologues. In line with this observation is that the *QCR8* gene from *S. pombe* could be cloned on the basis of complementation of an *S. cerevisiae* deletion strain [27], but not the *QCR7* gene from this organism. The requirements for subunit VII to allow assembly are apparently more strict than those for subunit VIII. Some of the homologues of subunit VII (from *P. falciparum* and *C. elegans*) were identified using the consensus sequence. It is likely that there are still several unidentified homologues of subunit VIII present in the databanks. The lack of a clear consensus sequence as a tool in screening the database would than be the reason for this.

The very low sequence conservation amongst the subunits VIII (see Fig. 4) indicates that proteins even less related will most likely lack sequence similarity. Subunit IV of the bc_1 complex from *Rhodobacter sphaeroides* has been suggested to fulfill the same function as the 9.5 kDa subunit from bovine heart, the subunit VIII homologue [28], this despite the fact that no amino acid sequence similarity could be found. However, structural features, such as hydrophobicity, are similar. The same features are conserved among the known subunit VIII homologues [27].

The SwissProt database searches that were performed using a profile of the multiple sequence alignment of both subunits did not find any protein showing similarity to the respective consensus sequences. For both subunits a role in the binding of ubiquinol/ubiquinone has been suggested [6–8]. It would be very useful if a consensus sequence for Q-binding could be defined. However, studies on other quinone binding proteins, such as cytochrome *b* [29] and those in the photosynthetic reaction center [30] indicate that the Q-binding pocket is formed by amino acids located far apart based on the primary structure or even in different proteins. We assume, therefore, that it is not possible to define a consensus sequence for Q-binding based on sequence alone.

The level of sequence conservation among the two protein families provides information about the function of these subunits within the bc_1 complex. It thus seems that subunit VIII predominantly requires certain structural features while subunit VII contains several specific amino acids required for the function of this protein.

In general we propose that the identification of homologous proteins and the definition of a consensus sequence provides targets for the design of mutant proteins for study of structure/ function relationships.

References

[1] Trumpower, B.L. (1990) Microbiol. Rev. 54, 101–129.
[2] Schägger, H., Link, T.A., Engel, W.D. and Von Jagow, G. (1986) Methods Enzymol. 126, 224–237.
[3] De Vries, S. and Marres, C.A. (1987) Biochim. Biophys. Acta 895, 205–239.
[4] Brandt, U., Uribe, S., Schägger, H. and Trumpower, B.L. (1994) J. Biol. Chem. 269, 12947–12953.
[5] Berden, J.A., Schoppink, P.J. and Grivell, L.A. (1988) in The Molecular Basis of Biomembrane Transport (Palmieri, F. and Quagliariello, E. eds.) pp. 195–208, Elsevier, Amsterdam.
[6] Hemrika, W., Berden, J.A. and Grivell, L.A. (1993) Eur. J. Biochem. 215, 601–609.
[7] Usui, S., Yu, L. and Yu, C.-A. (1990) Biochemistry 29, 4618–4626.
[8] Japa, S., Zhu, Q.-S. and Beattie, D.S. (1987) J. Biol. Chem. 262, 5441–5444.
[9] Schoppink, P.J., Berden, J.A. and Grivell, L.A. (1989) Eur. J. Biochem. 181, 475–483.
[10] Maarse, A.C., de Haan, M., Schoppink, P.J. Berden, J.A. and Grivell, L.A. (1988) Eur. J. Biochem. 172, 179–184.
[11] Klebe, R. J., Harries, J. V., Sharp, D. and Lacroute, F. (1980) Gene 11, 11–19.
[12] Soni, R. and Murray, J.A.H. (1992) Nucleic Acids Res. 20, 5852.
[13] Gietz, R.D. and Sugino, A. (1988) Gene 74, 527–534.
[14] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[15] De Haan, M., Van Loon, A.P.G.M., Kreike, J., Vaessen, R.T.M.J. and Grivell, L.A. (1984) Eur. J. Biochem. 138, 169–177
[16] Mulder, W., Scholten, I.H.J.M., Van Roon, H. and Grivell, L.A. (1994) Biochim. Biophys. Acta 1219, 719–723.
[17] Suzuki, H., Hosokawa, Y., Toda, H., Nishikimi, M. and Ozawa, T. (1988) Biochem. Biophys. Res. Commun. 156, 987–994.
[18] Wakabayashi, S., Takao, T., Shimonishi, Y., Kuramitsu, S., Matsubara, H., Wang, T., Zhang, Z. and King, T.E. (1985) J. Biol. Chem. 260, 337–343.
[19] Braun, H.-P., Kruft, V. and Schmitz, U.K. (1994) Planta 193, 99–106.
[20] Braun, H.-P. (1993) Ph.D. Thesis, Freie Universität, Berlin.
[21] Feng, D.F. and Doolittle, R.F. (1987) J. Mol. Evol. 35, 351–360.
[22] Schuler, G.D., Altschul, S.F. and Lipman, D. (1990) Proteins 9, 4355–4358.
[23] Gribskov, M., Lüthy, R. and Eisenberg, D. (1990) Methods Enzymol. 183, 146–159.
[24] Von Heijne, G. (1986) EMBO J. 5, 1335–1342.
[25] Hemrika, W., De Jong, M., Berden, J.A. and Grivell, L.A. (1994) Eur. J. Biochem. 220, 569–576.
[26] Huang, J.H., Struck, F., Matzinger, D.F. and Levings III, C.S. (1994) Plant Cell 6, 439–448.
[27] Boumans, H., Berden, J.A. and Grivell, L.A. (1995) Biochim. Biophys. Acta (in press).
[28] Usui, S. and Yu, L. (1991) J. Biol. Chem. 266, 15644–15649.
[29] Gennis, R.B., Barquera, B., Hacker, B., Van Doren, S.R., Arnaud, S., Crofts, A.R., Davidson, E., Gray, K.A. and Daldal, F. (1993) J. Bioenerg. Biomembr. 25, 195–209.
[30] Deisenhofer, J. and Michel, H. (1989) Science 245, 1463–1473.
[31] Maarse, A.C., de Haan, M., Schoppink, P.J. Berden, J.A. and Grivell, L.A. (1988) Eur. J. Biochem. 172, 179–184.
[32] Mulder, W., Scholten, I.H.J.M., Nagelkerken, B. and Grivell, L.A. (1994) Biochim. Biophys. Acta 1219, 713–718.
[33] Borchart, U., Machleidt, W., Schägger, H., Link, T. A. and Von Jagow, G. (1986) FEBS Lett. 200, 81–86.
[34] Braun, H.-P., Kruft, V. and Schmitz, U.K. (1994) Biochim. Biophys. Acta 1188, 367–372.