

# Close genetic relationship between *Nitrobacter hamburgensis* nitrite oxidoreductase and *Escherichia coli* nitrate reductases

# Karin Kirstein, Eberhard Bock

Institut für Allgemeine Botanik. Abteilung Mikrobiologie, Universität Hamburg, Ohnhorststrasse 18. D-22609 Hamburg, Germany

Received: 13 May 1993/Accepted: 5 July 1993

Abstract. The nitrite oxidoreductase (NOR) from the facultative nitrite-oxidizing bacterium Nitrobacter hamburgensis X14 was investigated genetically. In order to develop a probe for the gene norB, the N-terminal amino acid sequence of the NOR  $\beta$ -subunit (NorB) was determined. Based on that amino acid sequence, an oligonucleotide was derived that was used for the identification and cloning of gene norB. Sequence analysis of DNA fragments revealed three adjacent open reading frames in the order norA, norX, norB. The DNA sequences of norX and norB represented complete genes while the open reading frame of *norA* was truncated by the cloning site. The deduced amino acid sequence of protein NorB contained four cysteine clusters with striking homology to those of iron-sulfur centers of bacterial ferredoxins. NorB shares significant sequence similarity to the  $\beta$ subunits (NarH, NarY) of the two dissimilatory nitrate reductases (NRA, NRZ) of Escherichia coli. Additionally, the derived amino acid sequence of the truncated open reading frame of *norA* showed striking resemblance to the  $\alpha$ -subunits (NarG, NarZ) of the *E. coli* nitrate reductases.

Key words: Nitrobacter hamburgensis – Nitrite oxidoreductase genes – Nitrite oxidoreductase  $\beta$ -subunit – Ironsulfur protein – Escherichia coli nitrate reductase

The facultative lithotroph Nitrobacter hamburgensis derives energy for  $CO_2$  fixation via the Calvin cycle from the oxidation of nitrite to nitrate. In the absence of oxygen the cells are able to grow by nitrate reduction using organic material (Freitag and Bock 1990). Both, nitrite oxidation and nitrate reduction are catalysed by the membrane-bound nitrite oxidoreductase (NOR) (Tanaka et al. 1983; Sundermeyer-Klinger et al. 1984). The catalytically active enzyme consisted of two subunits with molecular weights of 115000 and 65000 when purified

by heat treatment (Meincke et al. 1992). In the presence of detergents cytochromes  $a_1$  and  $c_1$  were attributed to the NOR (Tanaka et al. 1983; Sundermeyer-Klinger et al. 1984). Cytochrome  $c_1$  was shown to be an integral membrane protein of a molecular weight of 32000 and considered to be the third subunit of the NOR (Sundermeyer-Klinger et al. 1984). Cofactors of NOR are molybdopterin (Krüger et al. 1987) and iron-sulfur centers (Meincke et al. 1992). The subunit composition of NOR and the ability to reduce nitrate are typical for dissimilatory nitrate reductases. As reviewed by Hochstein and Tomlinson (1988), dissimilatory nitrate reductases are membrane-bound, molybdenum-containing iron-sulfur proteins.

Escherichia coli possesses two dissimilatory nitrate reductases, designated NRA and NRZ. Chaudhry and MacGregor (1983) showed that the molybdenum cofactor is part of the  $\alpha$ -subunit while the  $\beta$ -subunit is supposed to be an electron-channeling Fe-S protein (Blasco et al. 1989, 1990). The cytochrome b containing  $\gamma$ -subunit is a membrane-embedded protein thought to be a link between the nitrate reductase and the respiratory chain.

Here we report on the identification, cloning and sequencing of the *norAXB* genes. The genes *norA* and *norB* encode the  $\alpha$ - and  $\beta$ -subunit, respectively, of NOR from *N. hamburgensis* X14.

# Material and methods

# Bacterial strains, plasmids and growth conditions

All bacterial strains and plasmids used in this work are listed in Table 1. *Nitrobacter hamburgensis* X14 was grown mixotrophically as previously described (Bock et al. 1983). *Escherichia coli* K12 was used as a host strain for cloning of DNA fragments and propagation of plasmids. The DNA vectors for cloning and sequencing were pIBI30 and pIBI31 (International Biotechnology Inc., New Haven, Conn., USA), and the lambda promoter vector pCE30 (Elvin et al. 1990) was used for expression of *norB* in *E. coh*.

#### Amino-terminal analysis of NorB

Membranes of nitrifying cells of *N. hambtargensis* were isolated according to Milde and Bock (1984). The membrane proteins were separated by sodium dodecyl sulphate-polyacrylamide gel electroTable 1. Bacterial strains and plasmids

Strain or plasmid	Relevant characteristics	Source/Reference
Nitrobacter		Bock et al. 1983
hamburgensis X14		V 1 1 Deserve et al
Escherichia coli	recAl endAl gyrA90 Im	Yanisch-Perron et al.
K12 JM109	hsdR17 supE44 relA1	1985; Ausubel et al. 1989
	$\Delta (lac-pro-AB)/F' [traD36]$	
	$proA^+B^+$ lac $I^q$ lac $Z \Delta M15$ ]	
Plasmids		
pIBI30	Ap <sup>r</sup> , <i>lacPOZ</i> ', 2.93 kb, mcsII	Internat. Biotech.
	$(\hat{E}coRI \gg HindIII)$	New Haven, Conn. USA
nIBI31	Ap <sup>r</sup> , <i>lacPOZ</i> ', 2.93 kb, mcsII	Internat. Biotech.
Parrie	$(HindIII \gg EcoRI)$	
nCE30	Ap <sup>r</sup> , Lambda c/857, lambda-p <sub>v</sub> lambda-p <sub>t</sub>	Elvin et al. 1990
pKK1	nIBI30: 19-kh EcoRI/BomHI	This study
picici	fragment from N hamburgensis	
	X14 currying part of norR	
~VV)	pIRI20: 7 kb EcoRI from	This study
prr2	M how how and a X14 comming	This study
	N. namourgensis A14, carrying	
	nor X B genes	TTI to standar
pKK2d	pKK2; deletion of 1.28-kb Xhol	This study
	fragment	
pKK3	pCE30; 2.7-kb SmaI fragment	This study
	from pKK2	

Ap<sup>r</sup>, resistant to ampicillin; mcs, multiple cloning site

phoresis (SDS-PAGE) according to Laemmli (1970). After electrophoresis, proteins were electroblotted onto polyvinyliden difluoride (PVDF) membrane as described by Matsudaira (1987). Then the PVDF membrane was stained with Coomassie blue and the membrane area carrying the NOR  $\beta$ -subunit was cut out and directly loaded into a gas-liquid phase sequenator (Model 470 A, Applied Biosystems; Funkstadt, Germany). The amino-terminal sequence of the NOR  $\beta$ -subunit was determined by automated Edman degradation (Edman 1956).

#### Recombinant DNA work

For routine work with recombinant DNA, established protocols were used (Sambrook et al. 1989). Chromosomal DNA of *Nitrobacter* was isolated using the method of Marmur (1961) modified by Koops and Harms (1985). The isolation of plasmids of *E. coli* was performed as described by Birnboim and Doly (1979).

#### DNA hybridization with an oligonucleotide probe

Chromosomal N. hamburgensis DNA or recombinant plasmids from E. coli were separated on an agarose gel and transferred to a Gene Screen Plus membrane (NEN Du Pont, Dreieich, Germany) by Southern blotting. Prehybridization was done for 30 min at 60 °C in a solution of 1 M NaCl, 1% (w/v) SDS, and 10% (w/v) dextran sulfate (50 µl/cm<sup>2</sup> filter area). Oligonucleotide synthesis was performed by means of a DNA/RNA synthesizer (Model 394. Applied Biosystems). The oligonucleotide probe was labeled with  $[\gamma^{-32}P]ATP$  using T4 polynucleotide kinase and then added to the prehybridized filter at a concentration of 10<sup>6</sup> cpm per ml of prehybridization solution. After an incubation for 18 h at 40 °C, the filter was washed twice for 5 min at room temperature in 0.3 M NaCl plus 30 mM sodium citrate  $(2 \times SSC)$ , twice for 30 min at 35 °C in  $2 \times SSC$ , 1% (w/v) SDS, and twice for 5 min at room temperature in  $2 \times SSC$  without SDS. The filter was wrapped in plastic foil and exposed to X-ray film.

#### DNA sequence analysis

The DNA sequence was determined by the dideoxynucleotide chain termination method of Sanger et al. (1977) according to the sequencing strategy shown in Fig. 2.

Nested deletions were generated by exonuclease III (Erase-a-Base Kit; Promega, distributed by Serva, Heidelberg, Germany). Sequencing kits (Pharmacia Sequencing Kit; Pharmacia, Freiburg, Germany; USB Taquence Version 2.0 DNA Sequencing Kit; United States Biochemical Corporation, Dannstadt-Schauernheim, Germany) were employed using  $[\alpha^{-35}S]$ ATP from Amersham-Buchler (Braunschweig, Germany).

#### Expression of norB in Escherichia coli

The expression of *norB* in *E. coli* was performed according to Elvin et al. (1990). *E. coli* JM109 containing the recombinant expression vector was grown at 30 °C in Luria-Bertani medium (Sambrook et al. 1989) containing ampicillin (50 µg/ml) to an OD<sub>595</sub> = 0.5, then treated at 42 °C for 2 h. Cells in 1-ml samples were harvested before and after heating, resuspended in an SDS-PAGE loading mixture at an OD<sub>595</sub> of 5. The samples were treated at 100 °C for 2 min prior to application of 20-µl samples to lanes of a 0.1% (w/v) SDS/10% (w/v) acrylamide slab gel. Following electrophoresis, proteins were stained with Coomassic brilliant blue.

#### Computer analysis

DNA and protein sequences were analysed by using the PC GENE (Genofit, Geneva, Switzerland) software package. Data search for sequence homology was performed by the PATMAT program (Henikoff et al. 1990) and sequence information was received from the EMBL Data Library (Heidelberg, Germany). Alignment of amino acid sequences was done by the MACAW program (Schuler et al. 1991).

#### Accession number

The sequence of the gene *norB* has been deposited in the EMBL Data Library under the accession number X66067 NHNORB.

Met1	Asp <sup>2</sup>	Ile <sup>3</sup>	Arg <sup>4</sup>	Ala <sup>5</sup>	Gln <sup>6</sup>	Val <sup>7</sup>	Ser <sup>8</sup>	Met <sup>9</sup>	Val <sup>1</sup>	<sup>0</sup> Phe <sup>1</sup>	<sup>1</sup> His <sup>1</sup>	2 <sub>Leu</sub> 1	<sup>3</sup> Asp <sup>1</sup>	<sup>4</sup> Lys <sup>15</sup>
5'-ATG	GAT C	ATC	CGT C	000 0	CAG	GTG	TCC G	ATG	GTG	110	CAC	CTG	GAT C	AA -3'
				Т			AG							

Fig. 1. N-terminal amino acid sequence of the NOR  $\beta$ -subunit (NorB) from *Nitrobacter humburgensis* X14 and the deduced sequences of the *norB*-specific 44mer oligodeoxynucleotides

#### Results

# The norB gene probe

The  $\beta$ -subunit (NorB) of the Nitrobacter hamburgensis NOR was isolated from purified membranes (Milde and Bock 1984) by using SDS-PAGE. The separated NorB was electroblotted onto PVDF membrane as described by Matsudaira (1987) and then directly loaded into an amino acid sequenator (Model 470 A, Applied Biosystems). The N-terminal amino acid sequence of NorB is shown in Fig. 1. Based on the codon usage of the N. vulgaris T3 genes for ribulose-1,5-bisphosphate carboxylase/oxygenase (E. Sickinger, personal communication), degenerated 44mer oligonucleotides were synthesized. The oligonucleotides used are listed below the amino acids in Fig. 1. After 5'-end labeling with [ $\gamma$ -<sup>32</sup>P]ATP, the degenerated oligonucleotides served as a norB gene probe.

# Isolation and sequencing of norB

The norB gene probe hybridized to a 1.9-kb EcoRI/ BamHI fragment and to a 7-kb EcoRI fragment of chromosomal *N. hamburgensis* DNA. The 1.9-kb fragment was isolated and inserted into the plasmid pIBI30, resulting in plasmid pKK1. Using the same techniques, the 7-kb fragment was also inserted into pIBI30 to generate plasmid pKK2. The latter plasmid contained the complete *norB* together with adjacent genes named *norX* and *norA* (Fig. 2) as revealed by sequencing.

The sequences of norB and adjacent DNA regions were determined by the dideoxy chain termination method (Sanger et al. 1977). The nucleotide sequence and the derived amino acid sequences are shown in Fig. 3. Upstream of *norB* the complete open reading frame (ORF) of norX and the incomplete ORF of norA were detected. The ORF of norB was identified by comparing the deduced amino acid sequences of all possible reading frames with the known N-terminal amino acid sequence of the NorB protein. ORF norB has a length of 1539 bp (positions 967–2505) and a coding capacity of 513 amino acids. The predicted molecular weight of NorB is 57995. The ORF is preceded by a putative ribosome-binding site at an appropriate distance (positions 952-957). At a distance of 32 nucleotides downstream of the stop codon TAG (positions 2506–2508) of *norB*, a region marked by arrows was found where dyad symmetry could form a



**Fig. 2.** Physical map of the partially sequenced 7-kb EcoRI fragment from *N. hamburgensis* X14 and sequencing strategy. The *arrows* show the positions, directions and the extent of sequences obtained. Subclones were produced by digestion of the plasmid pKK1 with different restriction enzymes. The subcloned fragments of pKK1 are depicted by *black bars*. The sequences obtained with these subclones

are indicated by the arrows 1-12. For sequencing the second half of *norB*, deletion plasmids of a derivative of plasmid pKK2 (plasmid pKK 2d) were constructed using exonuclease III. These regions are marked by *arrows* I-V. Based on the obtained sequence information from subclones synthetic oligonucleotides were designed and used as sequencing primers. These regions are indicated by the *arrows* a-1

450

10-bp stemmed hairpin loop. This region may represent a potential transcription termination structure whose free energy was calculated to be -97.1 kJ/mol.

The incomplete ORF of *norA*, ending with a TAG stop codon at positions 232-234, encloses 231 bp corresponding to a peptide of 77 amino acids. ORF *norX* extends for 648 bp (positions 250-898) and terminates with a TAA stop codon at positions 899-901. The gene *norX* may code for a polypeptide of 216 amino acids with a molecular weight of 23645. A putative ribosomebinding site (positions 233-238) is preceding the *norX* start codon at a distance of 11 bp.

No putative promoter sequence could be detected. The absence of a conspicuous promoter sequence upstream of norX or norB and the existence of a single putative termination structure downstream of norB indicate that

ORF norA=>

OKI	г л.	07A	=>						0.0.1		-	aaif		<b>0 1</b> m	-				CONC	60
GGC G	T	C	M	Y	Y	H	A	V	E	R	TAC	V	GTA Y	I	P	X	S	Q	E	60
CGT. R	AAC K	STGO W	ece:	rcc G	CGG G	CGG G	АСА Н	CAA N	CAG	CCT L	GAC T	GCG R	TAC	TCG R	TAT I	CAA N	ccc P	GCT L	GTTC F	120
CTG L	GCI A	COC:	G	CTA Y	TGC A	CCA Q	GTI F	CAC T	GTA Y	CGG G	CTG W	GAA N	CTA Y	TTG W	GGG G	P	GAC T	TGG G	TATC I	180
						~~~				~		~~~	~~~	~~~	~~~		0.003			240
L	T	R	D	T	H	V	v	v	R	K	M	E	K	L	E	W	**	*	Un IC	240
TCG	ca	LCGI	CATC	2470	CAT	GAT	2	GAR	CAT	GAT	CAT	GAA	cce	GCT	ccg	GCG	CAT	ттс	GGTG	300
			м	1	ж	I	Т	N	м	I	м	N	R	L	R	R	I	\$	۷	17
coc	ner	PEAR	COM	2077	тас	сат	caa	GAR	АТА	CGT	GAC	CAT	TAN	CTA	CAC	GCT	GCG	тас	CAAC	360
R	P	В	R	v	5	v	G	K	Y	v	r	Ľ	N	Y	T	L	R	Ť	N	37
aar	20	ገልሞ	cog	FCA	ĊAT	TCG	ATC	TGC	TOT	GTG	GGG	CGA	CGA	ACC	GTT	GGA	ата	TCA	GCAG	420
G	T	I	R	Н	I	R	5	A	v	W	G	D	E	₽	L	E	Y	Q	0	57
					~~m	<b>.</b>		00	000	~~~		oot	Table		cca	ecc	~~~	CGL	COTO	480
G	SS	G	R	L	L	P	A	Ľ	E	Ģ	K	A 001	A	G	E	R	7	E	v	77
												<b></b>								C 4 0
ATT	TT L	P	PCC	CGAL E	AGA E	GGC A	TTA Y	G	E	ACG R	AGA D	K	GGC À	L	Q	Q	R	V	P	97
GCT	GA	AAC	FT.	rcc	TGG	TGT	TGA	TCA	TAA	AGA	GCC	GGG	TAT	GIG	TTI	TCT	ççç	GCA	TAGC	600
A	E	Т	F	G	G	v	D	Q	I	Е	P	Ģ	M	¢	F	L,	A	н	5	117
GAC	GA	CGA	GCG:	rçg	GGT	TGA	GAA	CGT	GTT	GGT	CAC	CGA	GGT	TGA	TGA	GGA	TAA	CGG	GTAT	660
D	D	E	R	R	v	Е	N	v	L	۷	т	ε	۷	₽	Ë	D	N	G	Y	137
	0.00	-		<b></b>				ie of		~~~	олт	CAO	~~~	TO B		TO 3	cco	TLC	COTT	720
V	G TC V	V V	D	T	⇔,ALA N	H H	P	L	A.	G	M	T	L	E	F	E	G	s	v	157
	-	-	-	-																
GTG	GC;	AGT	3CGI	GGA	CAC	TCC	AAC	:000	GGC	GGG	ACA	TGT	CGA	TCC	GAC	GAT	GĞC	GAG	TATG	780
v	A	v	R	D	T	P	Ŧ	P	A	G	н	v	D	Р	T	л	А	3	n	±11
GGC	GA:	rgge	SCTV	GCC	GCT	TGÈ	TGC	GGC	GCA	AGC	GAC	AAT	TGA	сте	GCG	GCG	САА	GCG	GACG	840
G	D	G	L	A	L	A	A	A	Q	A	Т	I	D	s	R	R	к	R	т	197
n mat	~~~	BC 34	-1101				ътс		acc	cee	CAC	Tar		raa	Car	Tag	CAA	cac	атаа	900
I	D	Е	S	P	K	P	ŝ	Т	P	Ģ	Т	v	N	D	т	G	N	A	***	216
THE OF	ca:	5 3 CH	~na•	P & & (	n n n		ACG	AAG	COR	таа	GAA	CGA	AGC	CAT	AGG	AAC	GAA	GGA	GTAC	960
100	~~~																			
	no	7B=	>		<u></u>					~ ~ ~				~~~		<b>473 5</b>	<b>o</b> wo		CGGT	1020
TAC	AA	M	D	UAD I	R	AGO A	Q	.₩GT	S	M	V	F	H	L	D	ĸ	ĉ	I	G	19
TGC C	CA H	CAC T	CTG	CAG S	CAT I	CGC A	CTO: C	CA7 K	IGAA N	CAT	CIG	GAC T	CGA D	TCG R	CAA K	0000 C	CAC	CGA E	GTAC Y	1080
-		-	-	-	-		-													
ATG	TA	CTG	GAA	CAA	CGT	CGA	GAC	CAP	rečc	GGG	TAC	GGC	ATA	ccc	GAC	ACG	TIG	GGA	AGAC	1140
м	¥	W	ы	N	v	Е	т	ĸ	Р	G	т	G	r	2	T	ĸ	ж	Б	U	20
CAG	20	-	GTA	TCG	ccc	cco	CTO	GGT	CGT	CGA	CGG	GAC	GAG	GCA	GAP	GAG	cci	GCG	GCTG	1200
Q	T	K	Ŷ	R	G	G	W	v	۷	D	G	т	R	Q	к	3	L	R	г	78
										~ * *					~~~				-	1260
CGG	CT L	GCA	GGG <sup>.</sup> C	CRA. K	ATG W	000 C	AAC T	JGCI L	rgag S	⊂r⊾r N	IAL	CTI P	Y	CAA N	P	VI.N Y	L	P	T	98
~	-	ж				~	-	_	~		-	-	-		-	-	-			
CTC	GA	CGA	CTA	TTT	CGA	GCC	ATG	GAC	ста	CGA	CTP	TC	GAA	CCT	GAT	CAA	CGC	GCC	CCTG	1320
L	D	ס	Y	F	E	P	W	т	Y	D	Y	Q	М	L	I	N	A	P	L	118
GCC	GA	CGA	GCA	ACC	GAC	con	GCC	ccc	CAT	CTC	GAT	GGT	GAC	GGG	CAP	GTA	CAI	GGA	CACG	1380
	-		n	P	т	A	R	A	1	s	ж	v	Т	G	ĸ	¥	м	D	T	138

the genes *norA*, *norX*, and *norB* are probably organized in an operon.

The hydrophobicity profiles of NorB and NorX (not shown) suggested that both proteins are not membranebound.

# Expression of norB in E. coli

The calculated molecular weight of NorB is 57995. This value differs from that determined by SDS-PAGE (Sundermeyer-Klinger et al. 1984; Meincke et al. 1992). To prove that there have been deletions the gene *norB* was over-expressed in *Escherichia coli*. The gene product was compared with NorB from *N. hamburgensis* membranes.

For expression of *norB* the 2.7-kb SmaI fragment from pKK2 was inserted into the SmaI site of the lambda

I	CGA	GGO	GGG	CCC	gaa	CTG	GGA	ĊGA	CGA	TÇT	TGG	CGG	CIC	GCA	GGT	ста	CGC	CAA	CAAC	1440
	E	А	G	P	N	W	D	D	۵	L	G	G	S	ð	v	Y	A	N	N	158
GA	ree	GAA	СТТ	CGA	rcc	cec	стс	CGA	CGA	GGA	AAT	eca	CCA	GAT	CAA	CGA	GAT	CAA	CAGC	1500
D	P	N	F	۵	G	A	S	D	E	E	М	R	Q	Ĩ	N	E	I	N	S	178
													_					_		
AC <sup>i</sup>	GGT	CTT	TTT	CIA	CCT	SCC	GCG	CAI	CIG	CAA	CCA	TTG	TCT	CAA	TCC	GGG	CIG	CGT	SCCC	1560
T	۷	η.	4	I	-	r	R	+	6	n	Ħ	G	44	м	P	G	C	٠	A.	130
GÇI	GTG	ccc	GCA	GGG)	AGO	GAT	CTA	CAA	GCG	CGG	CGA	GGA	CGG	CGT	GGT	GCT	GGT	CAG	CCAG	1620
A	¢	2	Q	G	A	Ĩ	¥	ĸ	R	Ģ	Е	D	G	v	ν	L	v	5	Q	218
		-											~~~		~~ ~	~~~	~ . ~			1600
GAI P	506	CTG	CCG D	EGCI R	യ	10 10	URT M	GIG	UGT	<u>د ا</u> ت	2000 10	ore C	ື່ອ	v	CALA X	GAA R	GAC T	v	B DITC	238
-		Ŭ					••	-	•	-		Ť		-		-	-	-	-	
AA	CTG	GTC	GAC	CGGG	CAA	GGC	CGA	ллл	GTG	CAT	ССТ	GTG	CTA	TCC	GCG	CCT	CGA	GAG	CGGC	1740
N	W	s	т	G	ĸ	A	E	ĸ	С	I	L	С	Y	P	R	L	E	s	G	258
0.1	~~~~				-		ere	ста	com	cca	LCG	<b></b>	cca		ጥልጥ	raa	<b>r</b> rm	CGT	6¢me	1800
H	A	P	A	c	F	E	s	C	v	G	R	I	R	Ŷ	I	G	Ľ	v	L	278
TA	CGA	TGC	CGA	TGC	SAT	CGA	GGA	AAC	CGC	CAA	GGC	ccc	GCA	AGA	CCA	GTT	GGT	GAT	eece	1860
Y	D	A	D	A	Γ	Е	E	Ŧ	A	к	А	Ρ	Q	D	Q	Ľ	v	А	A	298
C24	202	<b>733</b>	CAT	- 14	-	1.CA	тео	GTT	таа	tee	TGA	ጥልጥ	CAT	cao	ddc.	тас	CAG	GĞĊ	CAAC	1920
õ	R	N	τ	Í	K	D	P	F	D	P	ם	I	Ι	A	A	A	R	A	N	318
-																				
GG'	TAT	ccc	CGA:	TTC	SAA(	JAT	CGA	GGC	eec	GCA	GAA	GTC	ecc	GGT	CTA	ÇĊA	GIT	CGT	CAAG	1980
G	I	P	D	5	к	I	E	A	A	đ	x	s	2	v	X	Q	μ.	¥	ĸ	118
AA	GTG	GGG	CAT	race	зсто	scc	GCT	GCA	TCC	GGA	GTT	ccg	CAC	GCT	GCC	GAT	GCT	GTT	CTAC	2040
x	w	G	I	A	L	F	L	H	P	E	F	R	т	L	₽	М	L	F	¥	358
GT	ACC	GCC	GCT	3GG(	3CC(	3GT	GCT	CGC	CAA	GGT	GGA	AAA	CGG a	CGT	GTA	CGA	CAA	CAT	TGCC	2100
¥.	5	-	-	G	5	٠	-	л	Ā	۷	-			v	•		14	-	-1	2,0
-								-	GAG	TTC.	ACT	CGA	AAG	GTC	GCG	TAT	ACG	сст	TOGO	2160
TU	GGA	ATC	GCG)	ACT	GGG:	rcc	$_{GTT}$	GWT												
rci S	GGA E	ATC S	GCG) R	ACT( L	3661 G	P P	GTT L	M	S	3	L	Е	R	s	R	I	R	L	R	398
TCI S	GGA E	ATC S	GCG R	L	G G	P	GTT L	M	S	S TCM	L	E	R	s	R	I COT	R	L	R	398
TCI S TA: Y	GGA E TAT	ATC S GGO A	GCG) R GAG S	L L CCT L	G G STT(	rcc P STC	GTT L GGG G	GRI M CGG G	S TAA N	S TGA E	L GGA D	E CAT I	R CAT I	S CCG R	R CGA- D	I CGT V	R CTA Y	L CAA K	R GAAG K	398 2220 418
TCI S TA: Y	GGA E TAT H	ATC S GGC A	GCG) R GAG S	L L L L	G G STT( L	rcc P STC S	GTT L GGG G	GAT M CGG G	S TAA N	S TGA E	L GGA D	E CAT I	R CAT I	S CCG R	R CGA D	I CGT V	R CTA Y	L CAA K	R GAAG K	398 2220 418
TCI S TA: Y CT/	GGA E TAT K K	ATC S GGC A CGC	GCG R GAG S GGT	ACT L DCT L SCG	G G STTC L CGT	TCC P STC S CTA	GTT L GGG G TAT	GAT CGG G GCG	S TAA N CTC	S TGA E GAA	L GGA D GAA	E CAT I CGT	R CAT I CAA	S CCG R GGA	R CGA D CAT	I CGT V TCO	R CTA Y GGA	L CAA K GGA	R GAAG K Agag	398 2220 418 2280
TO S TA: Y CT/ L	ega e tat k k agt v	ATC S GGC A CGC A	GCG R GAG S GGT V	ACTO L L SCGO R	G G STT( L CGT( V	TCC F STC S CTA	GTT L GGG G TAT M	GRT G G G R G CG R	S TAA N CTC S	S TGA E GAA R	L GGA D GAA K	E CAT I CGT V	R CAT I CAA K	S CCG R GGA D	R CGA D CAT I	I CGT V TCO P	R CTA Y GGA E	L CAA K GGA E	R GAAG K AGAG E	398 2220 418 2280 438
TCI S TA: Y CT: L CT:	EGA E TAT K AGT V	ATC S GGC A CGC A	GCG) R GAG S GGT V	ACT L CCT L SCG R	G G STT( L CGT) V	TCC F STC S CTA Y	GTT L GGG G TAT M AGG	GAT CGG G G CGG R	S TAA N CTC S GAC	S TGA E GAA R CAC	L GGA D GAA K	E CAT I GGT V CGC	R CAT I CAA K CGA	S CCG R GGA D GGT	R CGA D CAT I	I CGT V TCC P	R CTA Y GGA E GGT	L CAA K GGA E CTG	R GAAG K AGAG E GCGG	398 2220 418 2280 438 2340
TCI S TA: Y CT: L GT: V	GGA E TAT <sup>(</sup> M AGT <sup>(</sup> V SCAJ O	ATC S GGO A CGO A AAG R	GCG) R GAG S GGT V GGC A	ACT L CCT L SCG R SCT L	G G STT L CGT V SGC A	TCC STC S CTA CTA CGA	GTT L GGG G TAT M AGG G	GAT GGG GCG R CAA K	S TAA N CTC S GAC	S TGA E GAA R CAC T	L GGA D GAA K TGC A	E CAT I CGT V GGC	R CAT I CAA K CGA E	S CCG R GGA D CGT V	R CGA D CAT I CGA E	I CGT TCC P CGC	R TA Y GGA E GGT V	L CAA K GGA E CTG W	R GAAG K AGAG E GCGG R	398 2220 418 2280 438 2340 458
TCI S TA: Y CT: L GT: V	GGA E TAT <sup>I</sup> K AGT <sup>I</sup> V GCAJ Q	ATC S GGC A CGC A A AAG R	GCG R GAG S GGT V GGC A	ACT L CCT L SCG R SCT L	G G STT( L CGT( V CGC( A	CTA CTA CTA CGA E	GTT L GGG G TAT M AGG Q	GAT CGG G CGG R CAA K	S TAA N CTC S GAC T	S TGA E GAA R CAC T	L GGA D GAA K TGC A	E CAT I GGT V GGC A	R CAT I CAA K CGA E	S CCG R GGA D CGT V	R CGA- D CAT I CGA- E	I V TCC P GGC	R TA GGA E GGT V	L CAA K GGA E CTG W	R GAAG K Agag E GCGG R	398 2220 418 2280 438 2340 458
TCI S TA: Y CT/ L GTC V CT/	CGAI E TAT' M AGT' V GCAI Q CAC	ATC S GGC A CGC A A A A A CTC	GCG R GAG S GGT GGC A GTT	ACT L CCT L SCG R SCT L SCC	G G STT( L CGT( V CGC( A SAC)	TCC STC STC STC STC STC STC STC STC STC	GTT L GGG G TAT M AGG G CGA	GAI CGG G CAA K GGA	S TAA N CTC S GAC T ACG	3 TGA E GAA R CAC T	L GGA GAA K IGC A IGT	E CAT I CGT CC A GGT	R CAT I CAA K CGA E TCC	S CCG R GGA D CGT V GCC	R CGA D CAT I CGA E GAT	I CGT TCO P CGC A CGA	R TA GGA E GGT V GCG	L CAA K GGA E CTG W TGA	R GAAG K AGAG E GCGG R GACA	398 2220 418 2280 438 2340 458 2400
TCC S TA: Y CT/ L GTC V CT/ L CT/ L	GGAI E TAT M AGT V SCAI Q CAC T	ATC S GGC A CGC A AAG R CTC S	GCG R GAG S GGT V GGC A GTT L	ACT L CCT L CCG R CCG R CCT L CCC P	G G STT L CGT V CGC A SAC T	ICC F STC S CTA CCTA E CTT F	GTT GGG G TAT M AGG G CGA E	GAI CGG G CAA CAA K GGA	S TAA N CTC S GAC T ACG R	3 TGA E GAA K CAC T CTT F	L GGA D GAA K TGC A TGC V	E CAT I GGT GGC A GGT V	R CAT I CAA K CGA E TCC P	S CCG R GGA D SGC V SGCC P	R CGA D CAT I CGA E GAT M	I V TCC P GGC A GGA E	R Y GGA E GGT V GCG R	L CAA K GGA E CTG W TGA E	R GAAG K AGAG E GCGG R GACA T	398 2220 418 2280 438 2340 458 2400 478
TCC S TA: Y CTA: Y CTA: L GTC L GTC L GCC	GGA E TAT <sup>I</sup> M AGT <sup>I</sup> V SCAI Q CAC <sup>I</sup> T AGT <sup>I</sup>	ATC S GGCC A CGCC A AAG <sup>4</sup> R CTC S	GCG2 R GAG4 S GGT4 V GGC4 A GTT4 L TTC4	ACT L DCT L DCC R DCT SCC F DCC	G G G T C G C G C G C G C G C G C G C G	TCC F GTC S CTA C GA E CTT F	GTT L GGG G TAT M AGG G CGA E	GAT M GCGG R CAA K GGA E GCT	S TAA N CTC S GAC T ACG R CGA	3 TGA E GAA K CAC T T CTT F	L GGA D GAA K TGC A TGC V GGT	E CAT I GGT V GGC A GGT V GTC	R CAT I CAA K CGA CGA TCC P ACG	S GGA GGA CGT V GCCC P	R CGA D CAT I CGA E GAT M CTA	I CGT V TCC P GGC A GGA E CCC	R TA GGA E GGT V GCG R GGT	L CAA K GGA E CTG W TGA E TCG	R GAAG K Agag E GCGG R GACA T GAAG	398 2220 418 2280 438 2340 458 2400 478 2460
TCT S TA: Y CTT L CTT V CTT L GCL A	GGAI E TAT' M AGT' V GCAI CAC' T AGT' V	ATC S GGC A CGC A AAG R CTC S CGA D	GCG R GAG S GGC V GGC A GTT L S	ACTO L DOTO L SCCO R SCCO P SCCO P L	GGC: G STT L CGC V SGC A SAC STT F	ICC P GTC S CTA C CTA C C C C C C C C C C C C C C C	GTT GGGG G TAT M AGGG G CGA E CGA E	GAT GCGG GCG R CAA K GGA E GCT L	S TAA N CTC S GAC T ACG R CGA D	S TGA E GAA R CAC T CTT F TCO P	L GGA D GAA K TGC A TGT V CGT V	E CAT I GGT V GGC A GGC A GGC V GGC S	R CAT I CAA K CGA E TCC P ACG R	S CCG R GGA D SGCC P SCAA	R CGA- D CATT I CGA- E GAT- M CTA Y	I CGT V TCC P CGC A GGA E CCC P	R Y GGA E GGT V GCG R GGT V	L CAA K GGA E CTG W TGA E TCG R	R GAAG K Agag E GCGG R GACA T GAAG K	398 2220 418 2280 438 2340 458 2400 478 2460 498
TCT S TA: Y CT L GT V CT L GC. A	GGA E TAT M AGT V GCA Q CAC T AGT V	ATC S GGC A CGC A AAG R CTC S CGA D	GCGX R GAG S GGT V A GGC A S GTT L TTC S	ACT L DCT L SCG R SCT L SCC L SCC L	GGG G G G G G G G G G G G G T G G T G G T	ICC P S CTA C CTA C C C C C C C C C C C C C C C	GTT L GGGG G TAT M AGG G CGA E GCA Q	GGG G G CAA K GGA E GCT L	S TAA N CTC S GAC T ACG R CGA D	S TGA E GAA K CAC T CAC T F TCC P	L GGA D GAA K TGC A TGC V GGT V	E CAT I GGT V GGC A GGT V GTC S	R CAT I CAA K CGA E TCC P ACG R	S CCG R GGA D GGT V GGCC P CAA N	R CGA- D CAT I CGA- E GAT M CTA Y	I CGT V TCC P GGC A GGA E CCC P	R Y GGA E GGT V GCG R GGT V	L CAA K GGA E CTG W TGA E TCG R	R GAAG K Agag E GCGG R GACA T GAAG K	398 2220 418 2280 438 2340 458 2400 478 2460 498
TCT S TA: Y CTT L CTT L CTT L CTT L GCL A GC	GGA E TAT' K AGT' V SCAJ Q CAC' T AGT' V TGA'	ATC S GGC A CGC A AAG R CTC S CGA D GGT	GCG2 R GAG4 S GGT4 V GGC4 A GTT4 L TTC S CGG4	ACT L DCT L DCT L DCT L DCT L DCT L	GGG G G CGT <sup>I</sup> CGC <sup>I</sup> CGC <sup>I</sup> CGC <sup>I</sup> F CGC <sup>I</sup> CGC <sup>I</sup>	CCP GTC GTC CTA CCP CCP CCC CCC CCC CCC CCC CCC CCC CC	GTT L GGGG Q TAT M AGG G CGA CGA Q CCA	GGG GGG CAA K GGA E GGT L CAC	S TAAN N CTC S GACC T ACG R CGA D CGA	S TGA E GAA K CAC T T CAC T T CAC C CAC	L GGA D GAA K TGC A TGT V GGT V GGC	E CAT I GGT V GGC A GGC S GAG	R CAT I CAA K CGAA E TCCC P ACG R AGG	S GGA D GGCC P CAA N ACCC	R CGA D CAT I CGA E GAT M CTA Y ATA	I CGT V TCC P CGC A CCC P CCC P GGC	R T GGA E GGT V GCG R GGT V TTC	L CAA K GGA E CTG W TGA E TCG R TCG	R GAAG K AGAG GCGG R GACA T GAAG K TCGT	398 2220 418 2280 438 2340 458 2400 478 2460 498 2520 513
TCC S TA: Y CTI L GTC V CTC L GCC A GG <sup>o</sup> G	GGA E TAT' K AGT' V GCAJ CAC' T AGT V E	ATC S GGC A CGC A AAG R CTC S CGA D GGT V	GCG2 R GAG4 S GGC4 V GGC4 A GTT4 L TTC4 S G G	ACT L CCT L CCC R SCCC P SCCC P SCCC L CCT L	G G G C G C G C G G G G G G G G G G G G	ICC P GTC S CTA E CGA E CTT F CCC F	GTT L GGGG G TAT M AGGG G CGA CGA CGA CGA L CCA	GGG G G CAA K GGA E G CAC T CAC T	S TAAN N CTC S GACC T ACG R CGA D CGA D	S TGA E GAA K CAC T T CTT F TCC P CCCO P	L GGA D GAA K TGC A TGT V GGT V GGC A	E CAT I GGT V GGC A GGC A GGC S GAG R	R CAT I CAA K CGAA E TCC P ACG R ACG G	S CCCC R CCCC R D CCCC P CCAA N CCAA N CCAA P	R CGAA D CAT I CGA E GAT M CTA Y ATA **	I CGT V TCC P GGC A GGA E CCC P GGC	R Y GGA E GGT V GGG R GGT V TTC	L CAA K GGA E CTG W TGA E TCG R TTCG	R GAAG K AGAG E GCGG R GACA T GAAG K TCGT	398 2220 418 2280 438 2340 458 2400 478 2460 498 2520 513
TCC S TA: Y CTI L GTC V CTC L GCL A GCC CAC	GGA E TAT M AGT V SCA V SCA CAC T AGT V TGA CAC	ATC S GGC A CGC A AAG R CTC S CGA D GGT V TGC	GCG3 R GAG4 S GGT4 V GGC4 S GTT4 L TTC4 S CCG4 G AGG4	ACT L DCT L DCT R DCT R DCT P L DCT V CCC	G G G C G C G C G C G G G G G G G G G G	ICC P STC S CTA E CCA F CCC F CCC F ACA	GTT L GGGG G TAT M AGG G CGA CGA CCA E CCA H GCG	GCG GCG CAA K GGA E GCT T GTT	S TAANN CTC S GACC T ACG R CGA D CGA D	S TGA E GAA K CAC T T T C C C C C C C C C C C C C C C	L GGA D GAA K TGC A TGT V GGT V GGC A CAT	E CAT I CGT V GGCC A GGCC S GAG R ATĂ	R CAT I CAA K CGA K CGA TCC P ACG R ACG G CCC	S CCCG R GGAA D CCAA N CCAA N ACC P	R CGA- D CAT I CGA- E GAT- M CTA Y ATA- X ATA- CGC	I CGT P CGC A CCC P GGC CCC P GGC CCC CCC CCC CCC CCC	R Y GGA E GGT V GGCG R GGT V TTC GTC	L CAA K GGA E CTG W TGA TCG R TTCC TCG	R GAAG K AGAG E GCGG R GACA T GAAG K TCGT CGTA	396 2220 418 2280 438 2340 458 2400 478 2460 498 2520 513 2580
TCC S TA: Y CTI L GTC V CTI L GC. A GC. A GC. CA	EGA E TAT M M AGT V SCA C AGT V TGA E CGA	ATC S GGC A CGC A AAG R CTC S CGA D CTC S CGA V TGC	GCG2 R GAG4 S GGT4 V GGC4 A GTT4 L TTC0 S G G G A G G A G G	ACT L DCT L DCC R SCC R SCC P SCC V CCC V CCC	G G G C G C G C G C G G G G G G G G G G	ICC P STC S CTA E CTA F CCC F CCC F	GTT L GGGG G TAT M AGG G CGA CGA CCA H GCG	GCGG GCGG R CAAA E GGA E CACC T GTT	S TAA N CTC S GAC T ACG R CGA D CGA D GGG	S TGA E GAA K CAC T T TCO P CCC P	L GGA D GAA K TGC A TGT V GGT V GGC A CAT	E CAT I GGT V GGCC A GGCC S GAG R ATA	R CAT I CAA K CGA E TCCC P ACG R ACG G CCCC	S CCCG R GGAA D GCCC P CAAA N ACC P	R CGA- D CATI I CGA- E GAT- M CTA Y ATA- X **	I CGT V TCC P GGC C CCC P GGC C CCC P CGG	R Y GGA E GGT V GCG R GGT V TTC GTC	L CAA K GGA E CTG W TGA TCG R TCG	R GAAG K AGAG GCGG R GACA T GAAG K TCGT CGTA	396 2220 418 2280 438 2340 458 2400 478 2460 498 2520 513 2580
TCC S TA: Y CTI L GTC V CTC L GCC A GGC G GCA A TC	CGA E TAT N AGT V SCA Q CAC T TGA E CGA CGT	ATC S GGCC A CCCC A AAG R CTC S CCA CCCA TCC. CCCG	GCG2 R GAG4 S GGT4 V GGC4 G GT74 G GCT4 GCT4	ACT L DOT L DOT R SCC R SCC R SCC R SCC R SCC V C CC C CC	GGG G DGT CGC CGC CGC CGC CGC CGC CGC CGC CGC C	TCC P STC S S CTA S C S C C C C C C C C C C C C C C C C	GTT L GGGG Q TAT M GG G GCGA H GCG GCG	GAT GCGG GCG CAA K GGA E CAC T CAC T GTT CGA	S TAA N CTC S GAC T ACG R CGA D CGA D GGG GAG	S TGA' E GAA' K CAC' T T CCTT F CCCC P CCCC P TTT TTT.	L GGA D GAA K TGC A TGT V GGT V GGT V CAT	E CAT I GGT V GGC A GGC S GAG R ATA GAT	R CAT I CAA K CGA E TCC P ACG R ACG G CCC CCC	S CCG R GGA D GGCT P CCAA N ACC P AACC TCA	R CGA D CAT I CGA E GAT M CTA Y ATA CGC TCA	I CGT V TCC P GGC A GGA E CCC P GGC * CGG TAG	R TA Y GGA E GGT V GCG R GGT TTC GTC CAG	L CAA K GGA E CTG W TGA E TCG R TTCG R TTCG	R GAAG K AGAG E GCGG R GACA T GAAG K TCGT CGTA ATGA	396 2220 418 2280 438 2340 458 2400 478 2460 498 2460 498 2520 513 2580 2640
TCT S TA: Y CT L CT V CT L CT L C C C C C C C C C C C C C C C	CGAL E TAT' K AGT' V GCAL T CAC' T GAC T CAC' CGA' CGA' CGA' CGC'	ATC S GGC A CGC A AAG R CTC S CGA D CGA TGC CCG TTT	GCG2 R GAG4 S GGT4 V GGC4 R TTC S GGT4 S GCT4 S GCT4 TCC	ACT L ECT L ECG R SCT R SCT L SCC P L CGT V CGC CGC CAC	GGG G G CGT CGC A CGC A CGC G G G G G G G G G G G G	TCC P STC S CTA C CTA C C C C C C C C C C C C C C C	GTT L GGG G TAT M GGG CGA GCG GCG GGC GGC GGC GGC	GAT GCGG CAA K GGA E GCT CAC T GTT CGA GTT	S TAAA N CTC S GACC T ACG R CGA D CGA D GGG GGG GAG GAG	S TGA E GAA K CAC T T T C C T T T T T T T T T T T T T	L GGA D GAA K TGC A TGT V GGT V GGC A CAT ATA	E CAT I GGT V GGC A GGT S GAG R ATA GAT AAT	R CAT I CAA K CGA K CGC R ACG R ACG G CCCC CCCC CCCC CCCC	S CCCG R GGA D GCCT P CAAA N ACC P TCA	R CGA D CAT I CGA E GAT M CTA Y ATA ** CGC TCA	I CGT F CGT F CGC CGC CGC CGG CGG CGG TAG	R T GGA E GGT V GGG R GGT TTC CAG	L CAA K GGA E CTG W TGA TCG TCG TCG TCG TCG CCCC	R GAAG K AGAG E GCGG R GACA T GAAG K TCGT CGTA ATGA TTGT	398 2220 418 2280 438 2340 458 2400 478 2460 478 2520 513 2580 2640 2700
TCI S TA: Y CTI L GTC V CTI L GC. A GC. A GC. A GC. A TI GC. A TI CTI T CTI T CTI T CTI T CTI T CTI T CTI T CTI T CTI T CTI T CTI T CTI T CTI T CTI T CTI T CTI T CTI T CTI T CTI T CTI T CTI T CTI T CTI T CTI T CTI T CTI T CTI T C CTI T C CTI T C CTI T C CTI T C CTI T C CTI T C CTI T C CTI T C CTI T C CTI T C CTI C C C C	EGA E TAT' K AGT' V GCAC T AGT TGA E CGA CGT CGG	ATC S GGCC A CGCC A AAG R CTC S CGA T T CC CGA T T CC CGA	GCG3 R GAG4 S GGT4 V GGC4 S GGT4 S GGT4 S GGT4 S GGT4 S GGT4 S GGT4 S GGT4 S GGT4 S GGT4 S GGT4 S S GGT4 S S S S S S S S S S S S S S S S S S S	ACT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT L COT COT COT L COT L COT L COT L COT L COT L COT L COT L COT L COT COT COT COT COT COT COT COT COT COT	GGG GGT CGT CGC CGC CGC GG GG GG GG GG GG GG GG GG	TCC P GTC S CTA CGA CTT CCC F CCC F CCC F CCC F CCC F CCC F CCC F CCC F CCC F CCC F CCC F CCC F CCC F CCC F CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCC S CCCC S CCC S CCC S CCC S CCCC S CCCC S CCCC S CCCC S CCCC S CCCC S CCCC S CCCC S CCCC S CCCC S CCCC S CCCC S CCCC S CCCC S CCCC S CCCC S CCCC S CCCC S CCCC S CCCC S CCCC S CCCC S CCCC S CCCC S CCCC S CCCCC S CCCCCC	GTT L GGGG G TAT M GGA G GCA GCA GCG GCG GCG GCG GCG	GAT GCGG CAA CAA CAA CAA CAA CT CAA CT CAA CT CAA CT CAA CT CAA CT CAA CT CAA CT CAA CT CAA CT CAA CAA	S TAAN N CTC S GAC T ACG R CGA D CGA D GGG GGG GAG CTG	S TGA E CAC T CAC T T CO P C CO P T TT TAT	L GGA D GAA K TGC A TGT V GGT V GGC A CAT A TA	E CAT I GGT V GGC A GGT S GAG R AATÁ CAT	R CAT I CAA K CCA E TCC P ACG G CCC CCC CCC CCC CCC	S CCCG R GGA D GCCC P CCAA N ACCC P CCAA N ACCC TCA	R CGA D CAT I CGA E GAT M CTA Y ATA ** CGC TCA	I CGT P CGC A CCC P CGG CGG TAG GTT	R T GGA E GGT V GCG R GGT T T C AG T T C	L CAA K GGA E CTG W TGA TCG TCG TCG TCG ATC CCCC	R GAAG K AGAG E GCGG R GACA T GAAG K TCGT CGTA ATGA TTGT	398 2220 418 2280 438 2340 458 2400 478 2400 478 2520 513 2580 2640 2700 2736

Fig. 3. Nucleotide sequence of the *norB* gene and its flanking genomic DNA of *N. hamburgensis* X14 The proposed ribosome-binding sites are underlined and the deduced amino acid sequences of the ORFs *norA*. *norX*, and *norB* are indicated below the nucleotide sequence. *Asterisks* below the sequence identify stop codons. *Arrows* mark regions of dyad symmetry



Fig. 4. Comparison by SDS-PAGE of the NorB protein from Nitrobacter membranes and that overexpressed in Escherichia coli. Lanes B, NorB from Nitrobacter membranes (electroeluted from a gel of 2% (w/v) agarose in SDS-PAGE running buffer after electrophoresis of Nitrobacter membranes); pKK3, samples of E. colt containing pKK3 taken before  $(t_0)$  and 1  $(t_1)$  or 2 h  $(t_2)$  after exposure of the cells to 42 °C; pCE30, samples of E. colt containing pCE30 (without insert) taken before  $(t_0)$  and 1  $(t_1)$  and 2 h  $(t_2)$  after exposure to 42 °C; M, molecular weight markers: phosphorylase b (97400), bovine serum albumin (66200), ovalbumin (45000), carbonic anhydrase (31000), soybean trypsin inhibitor (21500) and  $\alpha$ -lactalbumin (14400) promoter vector pCE30 (Elvin et al. 1990) to produce pKK3. The protein patterns presented in Fig. 4 show that after the temperature shift from 30 °C to 42 °C the strain containing pKK3 overproduced a protein with an electrophoretic mobility very similar to the NorB protein isolated from N. hamburgensis membranes.

# Sequence comparisons

The predicted products of norA. norX, and norB were investigated by using data bank searches for sequence similarities. Proteins with significant sequence similarities were found for NorB and the incomplete NorA but not for NorX. NorB exhibited 45% and 46% sequence identity to the proteins NarH and NarY (Fig. 5) which are the  $\beta$ -subunits of the two dissimilatory nitrate reductases NRA and NRZ of E. coli (Blasco et al. 1989, 1990). Taking conservative amino acid changes into account, the calculated overall resemblance increases to 61% and 63%, respectively. The N-terminal region up to Lys-362 of NorB is characterized by especially high similarity. For instance, the four marked cysteine clusters are nearly identical in its amino acid sequence. The internal region (Gly-363 to Glu-477) is less conserved while the C-terminal section (Thr-478 to the C-terminus) exhibits significant variation.

The amino acid sequence derived from the incomplete gene *norA* shows significant similarity to the C-terminal sequences of the proteins NarG and NarZ (Fig. 6) which are the  $\alpha$ -subunits of the two dissimilatory nitrate reductases NRA and NRZ of *E. coli* (Blasco et al. 1989, 1990). Identities of 49% and 45.5% for NarG and NarZ,

NarH	MKIRSQVCHVLNLDKCIGCHTCSVTCKNVWTSREGVEYAWFNNVETKPGQGFPTDWENQEKYKGGWIRKINGKLQPRMGNRAMLLGKIFANPHLPGIDDYYEPFDFDYQNLHTAPEGS	118
NorB	MDIRAQVSMVFHLDKCIGCHTCSIACKNIWTDRKGTEYMYWNNVETKPGTGYPTRWEDQTKYRGGWVVDGTRQKSLRLRLQGKWGTLSNIFYNPYLPTLDDYFEPWTYDYQNLINAPLA-	119
Nar¥	MXIRSQVGMVLNLDKCIGCHTCSVTCKNVWTGREGMEYAWFNNVETKPGIGYPKNWEDQEEWQGGWVRDVNGXIRPRLGNXHGVITXIFANPVVPQIDDYYEPFTFDYEHLHSAPEG-	117
Narä	KSQPIARPRSLITGERMAKIEKGPNWEDDLGGEFDKLAKDKNFDNIQKAMYSQFENTFMMYLPRLCEHCLNPACVATCPSGAIYKREEDGIVLIDQDKCRGWRMCITGCPYKKIYFN	235
NorB	DEOPTARAISMVTGKYMDTIEAGPNWDDDLGGSQVYANNDPNFDGASDESMRQINEINSTVFFYLPRICNHCLNPGCVAACFOGAIYKRGEDGVVLVSQERCRAWRMCVSGCPYKKTYFN	239
Nary	KHIPTARPRSLIDGKRMDKVIWG9NWEELLGGEFEKRARDRNFEAMQKEMYGQFENTFMMYLPRLCEHCLNPSCVATCPSGAIYKREEDGIVLIDQDKCRGWRLCISGCFYKKIYFN	234
NarH	WKSGKSERCIFCYPRIEAGQPTVCSETCVGRIRYLGVLLYDADAIERAASTENEKDLYQRQLDVFLDPNDPKVIEQAIKDGIPLSVIEAAQQSPVYKMAMEWKLALPLHPEYRTLPMVWY	355
NorB	WSTGKAEKCILCYPRLESGHAPACFHSCVGRIRYIGLVLYDADAIEETAKAPODQ-LVHAQRNIIKDPFDPDIIAAARANGIPDSKIEAAQKSPVYQFVKKWGIALPLHPEFRTLPMLPY	358
NarY	WKSGKSEKCIFCYPRIESGQPTVCSETCVCRIRYLGVLLYDADRIEEAASTEREVDLYERQCEVFLDPHDPSVIEEALKQGIPQNVIDAAHGSPVYKMAMDWKLALPLHPEYRTLPMVWY	354
NarH	VPPLSPIQSAADAGELG-SNGILPDVESLRIPVQYLANLLTAGEYQTGTARTETYAGDASLQTAETVDGKVDTRALEEVGLTEAQAQEMYRYLAIANYEDRFVVPSSHREL	465
NorB	v pplgpvlakvengvydnlasesrlgplmsslersrirlrymasllsggnediirdvykklvavrvymrskkvkdipeeevgralaegkttaaeveavwrltslptpeerfvvppmeretingvykklvavrvymrskkvkdipeeevgralaegkttaaeveavwrltslptpeerfvvppmeretingvykklvavrvymrskkvkdipeeevgralaegkttaaeveavwrltslptpeerfvvppmeretingvykklvavrvymrskkvkdipeeevgralaegkttaaeveavwrltslptpeerfvvppmeretingvykklvavrvymrskkvkdipeeevgralaegkttaaeveavwrltslptpeerfvvppmeretingvykklvavrvymrskkvkdipeeevgralaegkttaaeveavwrltslptpeerfvvppmeretingvykklvavrvymrskkvkdipeeevgralaegkttaaeveavwrltslptpeerfvvppmeretingvykklvavrvymrskkvkdipeeevgralaegkttaaeveavwrltslptpeerfvvppmeretingvykklvavrvymrskkvkdipeeevgralaegkttaaeveavwrltslptpeerfvvppmeretingvykklvavrvymrskkvkdipeeevgralaegkttaaeveavwrltslptpeerfvvppmeretingvykklvavrvymrskkvkdipeeevgralaegkttaaeveavwrltslptpeerfvvppmeretingvykklvavrvymrskkvkdipeeevgralaegkttaaeveavwrltslptpeerfvvppmeretingvykklvavrvymrskkvkdipeevgralaegkttaaeveavwrltslptpeerfvvppmeretingvykklvavrvymrskkvkdipeevgralaegkttaaeveavwrltslptpeerfvvppmeretingvykklvavrvymrskkvkdipeevgralaegkttaaeveavwrltslptpeerfvvppmeretingvykklvavrvymrskkvkdipeevgralaegkttaaeveavwrltslptpeerfvvppmeretingvykklvavrvymrskvkdipeevgralaegkttaaeveavvrltslptpeerfvvppmeretingvykklvavrvymrskvkdipeevgralaegkttaaeveavvrltslptpeerfvvppmeretingvykklvavrvymrskvkdipeevgralaegkttaaeveavvrltslptpeevgralaegkttaaeveavvrltslptpeevgralaegkttaaeveavvrltslptpeevgralaegkttaaeveavvrltslptpeevgralaegkttaaeveavvrltslptpeevgralaegkttaaeveavvrltslptpeevgralaegkttaaeveavvrltslptpeevgralaegkttaaeveavvrltslptpeevgralaegkttaaeveavvrltslptpeevgralaegkttaaeveavvrltslptpeevgralaegkttaaeveavvrltslptpeevgralaegkttaaeveavvrltslptpeevgralaegkttaaeveavvrltslptpeevgralaegkttaaeveavvrltslptpeevgralaegkttaaeveavvrltslptpeevgralaegkttaaeveavvrltslptpeevgralaegkttaaeveavvrltslptpeevgralaegkttaaeveavvrltslptpeevgralaegkttaaeveavvrltslptpeevgralaegkttaaeveavvrltslptpeevgralaegkttaaeveavvrltslptpeevgralaegkttaaeveavvrltslptpeevgralaegkttaaeveavvrltslptpeevgralaegkttaaeveavvrltslptpeevgralaegkttaaeveavvrltslptpeevgralaegkttaaeveavvrltslptpeevgralaegkta	478
NarY	**** *: :: *: :: :: :: :: :: :: :: :: ::	465
NarH	areaffekngcgftfgdgchgsdtkfnlfnsrridaidvtsktephp 512 * : : ; * : ; *:	

NorB AVDSLFPQLDFVSRNYPVRKGEVGVGFHTDPARGP-----51.

Nary AGDAFAERNGCGFTFGDGCHGSDSKFNLFNSSRIDAINITEVRDKAEGE 514

Fig. 5. Alignment of the deduced NorB sequence to the NarH and NarY sequences of *E. coli*. The *asterisks* between the sequences indicate identical amino acids, and conservative amino acid changes (:) are defined as any within the following groups: I. L. M. V. A. G; D. E. N, Q; F. H, Y. W; K, R; S, T; C and P. Dashes indicate gaps introduced to optimize alignment. Cysteine clusters are *boxed* 

NarG	GMTMMYHAQERIVNLPGSEITQQRGGIHNSVTRITPKPTHMIGGYAHLAYGFNYYGTVGS-NRDBFVVVRKMKNIDWLDGEGNDQVQBSVI	1238
	* * *** ** * :* *: ***:*** * :::**** **:*** **:****	
NorA	gtcmyyhavertvyipksqerkwrggghnsltrtrinplflaggyaqftygwnywgptgiltrdthvvvrkmexlew	77
	* * *** ** : ** *: ****:*** * ::::******	
NarZ	GMTMMYHAQERIMNIPGSEVTCMRGGIHNSVTRVCPKPTHMIGGYAQLAWGFNYYGTVGS-NRDEFIMIRKMKNVNWLDDEGRDQVQBAKK	1246

Fig. 6. Alignment of the deduced sequence of the incomplete NorA peptide to the C-termini of NarG and NarZ of *E. coli*. For explanations see legend of Fig. 5

respectively, were calculated. Considering conservative amino acid changes, the values increase to 65% and 66%, respectively.

# Discussion

The genetic characterization of the key enzyme of Nitrobacter hamburgensis revealed that the  $\beta$ -subunit and part of the  $\alpha$ -subunit of NOR exhibit significant sequence similarities with the  $\beta$ - and  $\alpha$ -subunits of the two dissimilatory nitrate reductases of Escherichia coli. This represents the first genetic evidence for a close structural and functional relationship between NOR of N. hamburgensis and NRA and NRZ of E. coli.

The distribution of cysteines within the four NorB cysteine clusters is identical to the one in iron-sulfur centers of different bacterial ferredoxins as well as nitrate reductases and succinate dehydrogenase of *E. coli*. The amino acid composition of the cysteine clusters of NorB and of various electron transfer proteins is shown in Fig. 7. Cluster I contains a  $CX_2CX_2CX_3$  arrangement of the cysteines typical for the [4 Fe-4 S] centers of the ferredoxins I and II of *Desulfovibrio desulfuricans* Norway (Bruschi and Guerlesquin 1988). Cluster II,

 $CX_2CX_4CX_3C$ , is similar to the [3 Fe-3 S] center of Azotobacter vinelandii and Pseudomonas ovalis ferredoxins (Bruschi and Guerlesquin 1988). The arrangement of cysteines in Cluster III,  $CX_5CX_3C$ , is the same as described for Sulfolobus acidocaldarius and Thermoplasma acidophilum ferredoxins (Bruschi and Guerlesquin 1988). Furthermore, patterns of cluster III exhibit the same structure like that of the Fe-S subunit (SdhB) of succinate dehydrogenase of *E. coli*. The latter is supposed to form a [4 Fe-4 S] center (Darlison and Guest 1985). Finally, cluster IV,  $CX_2CX_{11}CX_3C$ , resembles the [4 Fe-4 S]-type found in the ferredoxins of Rhodopseudomonas palustris and Chromatium vinosum (Bruschi and Guerlesquin 1988).

The result of the comparison presented clearly shows that the four cysteine clusters of NorB share striking similarities with those of the  $\beta$ -subunits NarH and NarY of the two dissimilatory nitrate reductases NRA and NRZ of *E. coli* (Blasco et al. 1989, 1990). For NRA and NRZ three [4 Fe-4 S] clusters and one [3 Fe-4 S] cluster were demonstrated to be present using EPR spectroscopy (Johnson et al. 1985; Guigliarelli et al. 1992). Blasco et al. (1989) provided evidence for the existence of an Fe-S subunit that contained four iron-sulfur centers. Because of these findings, it seems reasonable to assume that NorB

		с		с		с				с								
N. hamburgensis	NorB (Cluster I)	сі	G	ĊF	I T	с	s	Ī	A	c :	ĸ							
E.coli	NarH (Cluster I)	сі	G	сF	łΤ	С	s	v	т	c :	ĸ							
E.coli	NarY (Cluster I)	сг	G	сE	Т	С	s	v	Т	c :	к							
Desulfovibrio desulfuricans Norway	FdI	С І	G	СE	S	С	v	ε	L	C	2							
D.desulfuricans Norway	FdII	сі	G	C G	E	С	v	D	V	C	P							
		с		с				с			¢	2						
N. hamburgensis	NorB (Cluster II)	CN	н	C L	J N	P	G	С	v	A	A C	P						
E. coli	NarH (Cluster II)	CE	H	c t	. N	P	A	С	v	A '	то	: P						
E. coli	NarY (Cluster II)	СЕ	H	¢ι	N	Р	δ	С	v	A	r c	P						
Azotobacter vinelandii	Fd	сі	K	сF	t Y	т	D	С	v	E '	v c	P						
Pseudomonas ovalis	Fd	СI	ĸ	C P	(Y	т	Þ	С	V	Ξ	vo	; P						
		c				с				C								
N. hamburgensis	NorB (Cluster III)	CR	A	WF	۲. M	C	v	s	G	c :	P							
E. coli	NarH (Cluster III)	CR	G 1	W R	٤м	C	I	т	G	c :	P							
E. coli	NarY (Cluster III)	CR	G	W F	۲Ľ	С	I	s	G	c :	P							
E. coli	SdhB (Cluster III)	СТ	F	V G	; Y	¢	s	Е	v	C	6							
Sulfolobus acidocaldarius	Fd	сI	A	DG	; s	С	I	т	A	c :	Р							
Thermoplasma acidophilum	Fd	CΙ	A I	DG	A	С	M	ם	V	C	P							
		с		с										с			с	
N homburgansis	Nor8 (Cluster IV)	- T	T.	r v	P	P	T.	8	5	2 1	нъ	P	A	С	Ŧ	H	S C	v
F coli	NarH (Cluster IV)	C T	F	c v	P	R	T	E	Ā	g (	5 F	· -	v	c	s	E	r c	v
E coli	NarY (Cluster IV)	C T	- F	cγ	P	R	Ť	E	s	G I	- τ ο τ	, <del>,</del>	v	č	s	Е	r c	v
Rhodonseudomonas nalustris	Fd	C T	Ē	с.				.8	~ a	а.				.c	v	A	v c	P
Chromatium vinosum	Fd	ст	E	c.				.8	a	a.,	• • •			. C	v	E :	r c	P

Fig. 7. Comparison of the cysteine clusters I–IV of NorB with those of various bacterial ferredoxins (Fd) as well as NarH, NarY and SdhB of *E. coli*. The Fd amino acid sequences are presented according to Bruschi and Guerlesquin (1988). The NarH and NarY amino acid sequences are from Blasco et al. (1989, 1990), and the SdhB amino acid sequence is from Darlison and Guest (1984)

may also contain three [4 Fe-4 S] centers and one [3 Fe-4 S] center. Therefore, NorB as a part of the membrane-associated protein complex of the NOR may function as an electron-channeling protein between the nitrite-oxidizing NorA, not yet characterized completely, and the membrane-integrated electron transport chain.

Even though the genetic analysis is not completed yet, the present results provide sufficient evidence for a close evolutionary relationship of the nitrate reductases from  $E.\ coli$  and the NOR from  $N.\ hamburgensis$ . This is consistent with the biochemical data which suggested close functional similarity between both enzyme complexes (Hochstein and Tomlinson 1988; Sundermeyer-Klinger et al. 1984). If this assumption is correct it might be speculated that NOR and nitrate reductases have a common ancestor. It is unknown whether the "ancient" nitrate reductase was able to catalyze the reverse reaction oxidizing nitrite to nitrate. To address this question further investigations on nitrite oxidoreductases and nitrate reductases should be performed.

Acknowledgements. This paper is based on a PhD thesis of Karin Kirstein submitted to the Faculty of Biology, University of Hamburg. The help of Dr. R. Bergmann in computer assisted sequence analysis is appreciated. The sequence analysis of NorB was done by Dr. F. Buck, and synthetic oligonucleotides were provided by Dr. W. Kullmann. c/o Prof. Dr. Richter, Universitätskrankenhaus Eppendorf, Institut für Zellbiochemie und Klinische Neurobiologie, Hamburg, Germany.

# References

- Ausubel FM, Brent R, Kingston RE, Moore DD, Smith JA, Seidman JG, Struhl K (eds) (1989) Current protocols in molecular biology. Wiley, New York
- Birnboim HC, Doly J (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res 7: 1513–1519
- Blasco F, Iobbi C, Giordano G, Chippaux M, Bonnefoy V (1989) Nitrate reductase of *Escheruchia coli*: completion of the nucleotide sequence of the *nar* operon and reassessment of the role of the  $\alpha$  and  $\beta$  subunits in iron binding and electron transfer. Mol Gen Genet 218: 249–256
- Blasco F, Jobbi C, Ratouchniak J, Bonnefoy V, Chippaux M (1990) Nitrate reductases of *Escherichia coli*: sequence of the second nitrate reductase and comparison with that encoded by the *narGHJI* operon. Mol Gen Genet 222: 104–111
- Bock E, Sundermeyer-Klinger H, Stackebrandt E (1983) New facultative lithoautotrophic nitrite-oxidizing bacteria. Arch Microbiol 136: 281–284
- Bruschi M. Guerlesquin F (1988) Structure, function and evolution of bacterial ferredoxins. FEMS Microbiol Rev 54: 155-175
- Chaudhry GR, McGregor CH (1983) Escherichia coli nitrate reductase subunit A: its role as the catalytic site and evidence for its modification. J Bacteriol 154: 387-394
- Darlison MG, Guest JR (1984) Nucleotide sequence encoding the iron-sulfur protein subunit of the succinate dehydrogenase of *Escherichia coli*. Biochem J 223: 507-517

- Edman P (1956) Mechanism of phenylisothiocyanate-hydantoin degradation of peptides. Nature 177: 667-668
- Elvin CM, Thompson PR, Argall ME, Hendry P, Stamford NPJ, Lilley PE (1990) Modified bacteriophage lambda promoter vectors for overproduction of protein in *Escherichia coli*. Gene 87: 123-126
- Freitag A, Bock E (1990) Energy conservation in Nitrobacter, FEMS Microbiol Lett 66: 157–162
- Guigliarelli B, Asso M, More C, Augier V, Blasco F, Pommier F, Giordano G, Bertrand P (1992) EPR and redox centers in nitrate reductases-A and reductases-Z from *Escherichia coli* – Evidence for a high-potential and low-potential class and their relevance in the electron-transfer mechanism. Eur J Biochem 207: 61-68
- Henikoff S, Wallace JC, Brown J (1990) Finding protein similarities with nucleotide sequence databases. Methods Enzymol 183: 111-132
- Hochstein LJ, Tomlinson GA (1988) The enzymes associated with denitrification. Annu Rev Microbiol 42: 231-261
- Johnson MK, Bennet DE, Morningstar JE (1985) The iron-sulfur cluster composition of *Escherichia coli* nitrate reductase. J Biol Chem 260: 5456-5463
- Koops HP, Harms H (1985) Deoxyribonucleic acid homologies among 96 strains of ammonia-oxidizing bacteria. Arch Microbiol 141: 214-218
- Krüger B, Meyer O, Nagel M, Andreesen JR, Meineke M, Bock E, Blümle S, Zumft WG (1987) Evidence for the presence of bactopterin in the cubacterial molybdoenzymes nicotinic acid dehydrogenase, nitrite oxidoreductase, and respiratory nitrate reductase. FEMS Microbiol Lett 48: 225-227
- Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of the bacteriophage T4. Nature 227: 680-685
- Marmur J (1961) A procedure for the isolation of deoxyribonucleic acid from microorganism. J Microbiol 3: 208-218
- Matsudaira I (1987) Sequence from picomole quantities of protein electroblottet onto polyvinyliden difluoride membranes. J Biol Chem 262: 10035-10038
- Meincke M, Bock E. Kastrau D, Kroneck PMH (1992) Nitrite oxidoreductase from *Nitrobacter hamburgensis*: redox centers and their catalytic role. Arch Microbiol 158: 127–131
- Milde K, Bock E (1984) Isolation and partial characterization of inner and outer membrane fractions of *Nitrohacter hamburgensis*. FEMS Microbiol Lett 21: 137-141
- Sambrook J, Fritsch EF, Maniatis T (eds) (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74: 5463-5467
- Schuler GD, Altschul SF, Lipman DJ (1991) A workbench for multiple alignment construction and analysis. Protein 9: 180-190
- Sundermeyer-Klinger II, Meyer W, Warninghoff B, Bock E (1984) Membrane-bound nitrite oxidoreductase of *Nitrobacter*: evidence for a nitrate reductase system. Arch Microbiol 140: 153-158
- Tanaka Y, Fukumori Y, Yamanaka T (1983) Purification of cytochrome  $a_1c_1$  from *Nitrobacter agilis* and characterization of nitrite oxidation system of the bacterium. Arch Microbiol 135: 265-271
- Yanisch-Perron C, Vieira J, Messing J (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of M13mp18 and pUC19 vectors. Gene 33: 103-119