Isolation and nucleotide sequence of the *Thiobacillus ferrooxidans* genes for the small and large subunits of ribulose 1,5-bisphosphate carboxylase/oxygenase

Victor Pulgar, Leonardo Gaete, Jorge Allende, Omar Orellana, Xavier Jordana and Eugenia Jedlicki

Departamento de Bioquimica, Facultad de Medicina, Universidad de Chile, Casilla 70086, Santiago 7, Chile

Received 3 September 1991

The genes encoding for the large (*rbcL*) and small (*rbcS*) subunits of ribulose-1.5-bisphosphate carboxylase (RuBisCO) were cloned from the obligate autotroph *Thiobacillus ferrooxidans*, a bacterium involved in the bioleaching of minerals. Nucleotide sequence analysis of the cloned DNA showed that the two coding regions are separated by a 30-bp intergenic region, the smallest described for the RuBisCO genes. The *rbcL* and *rbcS* genes encode polypeptides of 473 and 118 amino acids, respectively. Comparison of the nucleotide and amino acid sequences with those of the genes for *rbcL* and *rbcS* found in other species demonstrated that the *T. ferrooxidans* genes have the closest degree of identity with those of *Chromatium vinosum* and of *Alvinoconcha hessleri* endosymbiont. Both *T. ferrooxidans* enzyme subunits contain all the conserved amino acids that are known to participate in the catalytic process or in holoenzyme assembly.

CO2 fixation; Nucleotide sequence; Amino acid sequence; Thiobacillus ferrooxidans

1. INTRODUCTION

Thiobacillus ferrooxidans is an acidophilic, chemolithotrophic, Gram-negative bacterium that participates in bacterial leaching of minerals [1]. It is an obligate autotroph fixing atmospheric CO_2 via the Calvin cycle [2]. The key enzyme of this pathway, ribulose bisphosphate carboxylase-oxygenase (RuBisCO) from T. ferrooxidans has been isolated and characterized in our laboratory [3]. The enzyme has a native M_r of 570 000 and is composed of M_r 54 000 and M_r 15 500 subunits, corresponding to a type I enzyme with its characteristic hexadecameric structure, L_8S_8 . The enzyme has apparent K_m values for CO₂ and for ribulose bisphosphate of 28 μ M and 80 μ M. respectively, and it can be inhibited by 6-phosphogluconate. It was estimated that RuBisCO accounts for 2-5% of the total protein present in a whole extract, and its specific activity increases when the percentage of the CO_2 is raised in the air bubbling through the culture [3].

RubisCO is present in a wide range of autotrophic organisms, from prokaryotes to higher plants. The L_8S_8 structure for the holoenzyme is found in the majority of the bacteria, algae and higher plants studied. A less common form of RuBisCO which contains only large subunits, corresponding to the type II enzyme, is found in some purple non-sulfur photosynthetic bacteria such as *Rhodospirillum rubrum* [4]. *Rhodobacter sphaeroides* and *Rhodobacter capsulatus* synthesize both type I and type II RuBisCO [5,6].

Some of the genes encoding for the large subunit of this enzyme in different organisms have been sequenced, showing a high degree of sequence similarity [7-13]. However, a great heterogeneity in nucleotide sequences has been found in genes encoding for the small subunit [8,12–15].

The growth of *T. ferrooxidans*, an economically important microorganism, appears to be regulated by its capacity to fix CO_2 . For this reason, the study of the RuBisCO genes and the regulation of their expression has special biotechnological interest. Recently, Kusano et al. reported the cloning and expression of *T. ferrooxidans* RuBisCO in *E. coli*, confirming the hexadecameric form of the enzyme [16]. This group, however, did not publish the sequence of the genes which they cloned. In this communication we present the cloning and sequencing of *T. ferrooxidans* rbcL and rbcS genes.

2. MATERIALS AND METHODS

2.1. Bacterial strains and phages

E. coli JM105, maintained in minimal medium supplemented with glucose was used as a recipient for recombinant M13 DNA. *E. coli* C600 transformed with recombinant plasmids was grown in Luria broth supplemented with 100 μ g of ampicillin per ml of culture.

T. ferrooxidans strain ATCC 19859 was grown in Mackintosh medium [17]. All recombinant DNAs were constructed according to standard methods [18].

Correspondence address: E. Jedlicki, Departamento de Bioquimica, Facultad de Medicina, Universidad de Chile, Casilla 70086, Santiago 7, Chile, Fax: (56) (2) 376321.

2.2. Chromosomal DNA library

A T. ferroxidans strain ATCC 19859 gene library was constructed by total digestion of chromosomal DNA with BamHI and ligation of the resulting fragments to pBR322 previously linearized by digestion with BamHI and dephosphorylated by treatment with calf intestine alkaline phosphatase [18].

2.3. Screening of the DNA library

The nearly complete *rbcL* gene from *A. nidulans*, cloned in pCS751 plasmid (kindly provided by Dr F.R. Tabita [19]), was labelled with $[\alpha^{-32}P]$ dATP by nick translation as described [18] and was used as a probe to screen the *T. ferrooxidans* gene library by the method described by Hanahan et al. [20].

 10×10^3 recombinant colonies were screened by the colony hybridization method [20], and several colonies from the first screening were purified through secondary and tertiary screening.

Hybridizations were performed overnight at 42° C in a solution containing 5×SSPE (50 mM sodium phosphate, 5 mM EDTA, 0.9 M NaCl), 1×BFP (0.02% BSA, 0.02% Ficoll, 0.02% PVP), 20% formamide, 0.5% SDS and 25 µg per ml of heat-denatured salmon sperm DNA.

Filters were washed four times at 42° C in $4 \times$ SSPE and then twice at 52° C for 15 min, before exposing to autoradiographic films at -70° C for 2-3 days.

2.4. DNA sequencing

Appropiate restriction fragments from recombinant plasmids were subcloned into M13mp18 or M13mp19 vectors, according to the instructions from the manufacturer (Bethesda Research Laboratories). Nucleotide sequences were determined using the dideoxy chain termination method [21]. Sequencing reactions were carried out using the Sequenase Kit supplied by USB Corp. [α -³⁵S]dATP (Amersham) was used as the radiolabelled nucleotide. Synthetic oligonucleotides were used as primers.

3. RESULTS AND DISCUSSION

3.1. Screening of T. ferrooxidans chromosomal DNA library

The isolation and identification of the genes coding for RuBisCO was carried out using a *T. ferrooxidans* gene library and the *rbcL* gene from *A. nidulans*. Approximately 10^4 colonies were screened and 20 positive clones were obtained. Eight of these clones were purified by secondary and tertiary screening. DNA was prepared from three purified clones, digested with different restriction enzymes, blotted into nitrocellulose membranes and hybridized to the labeled *rbcL* gene from *A. nidulans*.

The restriction map of one recombinant plasmid, pRB-1, is presented in Fig. 1. Similar analysis of the other recombinants showed that the three purified clones were identical. The 3.3 kbp *Eco*RI-*Bam*HI fragment was subcloned in pBR322 (pRB-2) and some restriction sites were identified. DNA sequencing was performed on fragments derived from both recombinant plasmids, the sequencing strategy being summarized in the lower part of Fig. 1.

3.2. Nucleotide sequences of rbcL and rbcS genes

The complete *rbcL* and *rbcS* genes from *T. ferrooxidans* were present in the DNA fragment cloned in pRB-



Fig. 1. Physical maps and the strategy for sequence analysis of the *T. ferrooxidans rbcL* and *rbcS* genes. A physical map of the 6.6-kbp *Bam*HI fragment (pRB-1) and of the 3.3-kbp *Eco*RI-*Bam*HI fragment (pRB-2) are shown. The arrows indicate the extent and direction of sequences determined and the solid bars show the regions encoding *rbcL* and *rbcS*. The direction of transcription, and some restriction sites are presented. A, *Ava*I; AII, *Ava*II; B, *Bam*HI; E, *Eco*RI; H, *Hind*III; K, *Kpn*I, P, *Pst*I and S, *Sa*II.

1 and its entire nucleotide sequence as well as the deduced amino acid sequences are shown in Fig. 2. The 2209 bp sequence predicts two consecutive open reading frames. The large and small subunit sequences are 1422 and 357 nucleotides long, respectively, leaving an intergenic region of 30 bp, the smallest described for prokaryotic RuBisCO genes. In prokaryotes the two subunit genes are co-transcribed, while in higher plants the large subunit is encoded by chloroplast DNA and the small subunit is nuclear encoded [22].

The *rbcL* gene encodes for a protein of 473 amino acid residues while *rbcS* encodes for a polypeptide of 118 amino acid residues, with calculated molecular weights of 52.7 and 13.6 kDa, respectively.

Possible rbcL and rbcS ribosome-binding sequences at positions -12 to -7 and 1442 to 1446, respectively, were detected (underlined in Fig. 2), but no consensus sequences for transcription promoters were found in the cloned fragment.

Two potential stem-loop sequences were found adjacent to the 3' end of rbcS (arrows in Fig. 2), the last one being followed by the sequence TTTT (dotted line in Fig. 2), which could represent a rho-independent transcriptional termination signal.

3.3. Comparison of nucleotide and amino acid sequences

Table I shows the results of nucleotide and amino acid sequence comparisons of the RuBisCO genes from *T. ferrooxidans* with those of other bacteria such as *C.* vinosum [8], *A. hessleri* endosymbiont [12], *A. nidulans* [7,14], *A. eutrophus* [13] and *R. rubrum* [10]. rbcL and rbcS nucleotide sequences from *T. ferrooxidans* exhibit the highest similarity with *C. vinosum* and *A. hessleri* endosymbiont sequences.

Predicted *rbcL* and *rbcS* gene products from *T. ferro*oxidans also showed the highest percentage of similarity at amino acid level with *C. vinosum* and *A. hessleri* FEBS LETTERS

		-20 GAA	0 CGGC.	AGT	-190 ACCG	ATAT	- Ig A	180 GTTA	GTCC	-1 N TC	70 ATTA	GCAC	-16 CAG	0 CTITA	ACA	-150 CTTT	ATGC	- TT C	140 GCTC	GTTA	-1 G TT	30 GTGT	TGGA	-12 ATT	0 GTGA	GCG	-110 TAAA	CAAT	PT		
	-100			-90		-1	-80		-70		-60			-50			-40		-3	-30		-20			-10						
		CAC	ACAG	GAA	ACAG	TCTA	JA C	CTAG	ATTCO	C GC	AAGC	TTGC	TAG	TCCG	CAG	ACGC	AGCG	IC T	CCGG	GGCA	C CG	GCGT	CGCG	CAA	CTGC	TAG	GAGG	rege	AT		
ATG	GCC	GTC	AAG	ACC	TAT	AAC	GCG	GGT	GTG	AAG	GAT	TAC	CGG	AAC	ACC	TAT	TGG	GAA	CCG	GAT	TAC	AGC	GTC	AAG	GAT	ACC	GAT	ATC	CTT	GCC	GTG
MET	Ala	Val	Lys	The	TYP	Asn	ALA	GTÅ	vai	гув	Абр	Tyr	Arg	ABR	144	TYP	Trp	610	Pro	veb	Tyr	Ser	Val	Lys	Авр	Thr	Авр	110	Lou	A18	Va1 192
Phe	AAG Lys	ATT Ile	ACT Thr	Pro	CAG Gln	GCG Ala	GGC Gly	GTG Val	GAT Asp	CGT Arg	GAG Glu	GAG Glu	GCG Ala	CCT Pro	GCT Ala	GCG Ala	GTT Val	GCC Ala	GCC Ala	GAA Glu	TCC Ser	TCT Ser	ACC Thr	GGC Gly	ACC Thr	TGG Trp	ACC Thr	ACG Thr	GTG Val	TGG Trp	ACG Thr
GAC	CTT	стс	ACC	GAC	CTT	GAC	TAC	TAC	AAG	GGG	CGC	GCA	тас	AGG	240 ATT	GAA	GAC	GTT	CCG	GGC	gat	GAT	ACC	TGT	TTC	TAT	GCG	ттс	ATT	GCC	288 TAC
Авр	Leu	ron	Thr	Asp	Leu	Asp	Tyr	Tyr	Lys	Gly	Arg	Ala	Τyr	Arg	11e 336	Glu	Asp	Val	Pro	Gly	Asp	Авр	Thr	Сув	Phe	Tyr	Ala	Phe	Ile	Ala	Tyr 384
CCC Pro	ATC Ile	GAT And	CTC Lou	TTC Phe	GAA Glu	GAA Glu	GGT Glv	TCC Ser	GTG Val	GTC Val	AAC Asn	GTC Val	TTC Phe	ACC Thr	TCG Ser	CTG Leu	GTT Val	GGT Glv	AAC Agn	GTC Val	TTC Pho	GGC Glv	TTC Pho	AAG Lvs	GCG Ala	GTA Val	CGT Arg	GCC Ala	CTG Leu	CGC Arg	CTG Leu
GAA	CAC	GTC	ССТ	-	CCC	አጥሮ	600	TAC	GTC	220	ACC	TGC	GGT	666	432	ccc	CAC	000	8 TTT	CAG	GTC	GAA	CGC	GAC	ATC	ATG	AAC	AAG	ፕልሞ	GGT	480
Glu	Asp	Val	Arg	Phe	Pro	Ile	Ala	Tyr	Val	Lys	Thr	Сув	Gly	Gly	Pro	Pro	His	Gly	Ile	Gln	Val	Glu	Arg	Asp	Ile	Mot	Asn	LYS	Tyr	Gly	Arg
ССА	TTG	CTG	GGC	TGC	ACC	ATC	AAA	ccc	AAG	CTG	GGT	TTG	TCG	GCC	AAG	AAC	TAT	GGT	CGC	GCT	TGC	TAT	GAG	GGT	CTG	CGT	GGT	GGA	стс	GAC	TTC
Pro	Leu	Leu	ClÀ	Сүз	Thr	Ile	Lys	Pro	Lys	Leu	Gly	Lou	Ser	Ala	Lys 624	Asn	Tyr	Gly	Arg	Ala	Сув	Tyr	Glu	GIÀ	Leu	Arg	GIY	GIΥ	Leu	Asp	Phe 672
ACT Thr	AAA Lys	GAT Asp	GAC Asp	GAA Glu	ААС Авл	GTC Val	аат Ава	AGC Sør	CAG Gln	CCT Pro	TTC Phe	ATG Møt	CGC Arg	TGG Trp	AGG Arg	CAG Gln	CGT Arg	TTC Phe	GAC Asp	TTC Phe	GTC Val	ATG Met	GAG Glu	GCC Ala	ATC Ile	CAG Gln	AAG Lys	GCC Ala	GÀG Glu	GCC Ala	GAG Glu
ACC	GGA	GAG	CGT	AAG	666	CAC	TAC	CTG	AAC	GTT	ACC	GCC	CCG	ACT	720 CCG	GAG	GAA	ATG	TAC	AAG	CGT	GCG	GAG	TAC	GCC	AAG	GAA	ATC	GGC	GCA	768 CCC
Thr	Gly	Glu	Arg	Lys	C1À	His	Tyr	Leu	Asn	Val	Thr	Ala	Pro	Thr	Pro 816	Glu	Glu	Het	Tyr	Lys	Arg	Ala	Glu	Tyr	Ala	Lys	Glu	Ile	Gly	Ala	Pro 864
ATC	ATC	ATG	CAC	GAT	TAC	ATC	ACC	GGC	GGC	TTC	TGC	GCC	AAC	ACG	GGT	CTG	GCC	AAC	TGG	TGC	CGC	GAC	AAC	GGC	ATG	CTC	CTG	CAC	ATT	CAC	CGC
176	110	Met	878	ASP	TYP	110	Thr	GIÀ	GIY	FUO	Сув	AIS	AUD	Thr	912	ren	AIA	АВП	Trp	Cys	Arg	Азр	4 81	GIY	MOL	Ter	760	018	119		960
GCC Ala	ATG Met	CAT His	CGG Arg	GTG Val	CTG Leu	GAC Авр	CGC Arg	ААС Авп	CCT Pro	CAC His	CAC His	GGC Gly	ATC Ile	CAC His	TTC Pho	CGT Arg	GTG Val	CTG Lou	ACC Thr	AAG Lyg	ATT Ile	CTC Leu	CGT Arg	CTG Leu	TCC Ser	GGC Gly	GGC Gly	GAT Абр	CAC Hig	Leu	CAC Bis
TCG	GGT	ACC	GTG	GTC	GGC	AAG	стс	GAA	GGC	GAC	CGT	GAG	GCG	1 ACC	CTG	GGC	TGG	ATC	GAC	ATC	ATG	CGT	GAT	AGA	TTC	АТС	лаg	GAA	GAT	CGC	.056 AGC
Ser	Gly	Thr	Val	Val	Gly	Lys	Leu	Glu	Gly	Авр	Arg	Glu	Ala	Thr 1	Leu 104	Gly	Trp	110	Asp	Ile	Met	λrg	Asp	Arg	Phe	110	Lys	Glu	Asp	Arg 1	Ser 1152
CGA	GGC	ATC	TTC	TTC	GAC	CAG	GAC	TGG Trp	GGC	TCC	ATG	CCG	GGT	GTC Val	ATG	CCG	GTC Val	GCT Ala	TCC Ser	GGT Glv	GGT Glv	ATC Ile	CAC His	GTT Val	TGG Trp	CAC His	ATG Met	CCG Pro	GCG Ala	CTG Leu	GTC Val
nu y			200	e	~~~		anc.		013		000		011	1	200		C M					COT	MG	GGT		000	600	200	CCC	I GTC	248
ACC Thr	ATC Ilo	Phe	GGT Gly	Asp	Asp	Sor	Val	Leu	Gln	Phe	Gly	Gly	Gly	Thr	Leu	Gly	His	Pro	Trp	Gly	Asn	Ala	Lys	Gly	Ala	Ala	Ala	Asn	Arg	Val	Ala
CTG	gaa	GCC	TGT	GTG	GAA	GCC	CGC	AAC	CGC	GGT	GTG	GCC	ATC	GAG	AAG	GAA	GGC	AAG	GCT	GTT	стс	ACT	GAG	GCG	GCC	AAG	CAC	тсс	ccc	GAA	CTC
Leu	Glu	Ala	Сув	Val	Glu	Ala	Arg	Asn	Arg	Gly	Val	Ala	Ile	Glu 1	LY8 392	Glu	Gly	Lys	Ala	Val	Leu	Thr	Glu	Ala J	A1a 422	Lys	H10	ser	FIO	GIU	ren
AAG Lvs	ATT Ile	GCC Ala	ATG Met	GAA Glu	ACG Thr	TGG Trd	AAA Lvs	GAG Glu	ATC Ile	AAG Lvs	TCC Ser	GAA Glu	TTC Phe	GAT Asp	ACG Thr	GTC Val	GAC Asp	AAG Lys	CTG Leu	GAC Авр	GTT Val	GCC Ala	CAT His	AAA Lys	TAA						
-2-		GCAC	14	132 AT 0	CACCI	144	2 A GG	AGTO	1452					•			•	-													
		den.								CTC	mee	GAC		1	500	666	AAG	ማምሮ	GAG	ACA	CTG	TOT	TAC	CTG	ccc	GCC	CTG	ACC	GCG	1 GAC	548 GAG
ATG MET	Ser	Glu	Val	Gln	Asp	Tyr	лял Lys	Ser	Arg	Leu	Ser	Авр	Pro	Ala	Ser	Ala	Lys	Phe	Glu	Thr	Leu	Ser	Tyr	Leu	Pro	Ala	Leu	Thr	Ala	Asp,	Glu 644
ATC	CGT	CAA	CAG	GTT	GCG	TAT	ATT	GTT	тсс	AAG	GGC	TGG	AAT	CCG	GCG	GTA	GAA	CAT	ACC	GAA	CCG	GAA	AAC	GCC	TTC	GGC	AAC	TAC	TGG	TAT	ATG
Ile	Arg	Gln	Gln	Val	Ala	Tyr	Ile	Val	Ser	Lys	Gly	Trp	Asn	Pro. 1	Ala 692	Val	Glu	His	Thr	Glu	Pro	Glu	Asn	Ala	Pue	GΙΫ	ABR	uγr	TIP	1	740
TGG	AAG Lys	TTG Leu	CCC Pro	ATG Met	TTC Phe	GGC Glv	GAA Glu	ACG Thr	GAC Asp	GTG Val	GAC Asp	ACC Thr	ATT Ile	CTG . Leu	AAA Lyg	GAA Glu	GCG Ala	GAA Glu	CGC Arg	TGC Cys	CAT Bis	AAG Lys	CGG Arg	AAT Asn	CCC Pro	CAT Hio	AAC Asn	CAC His	GTC Val	CGT . Arg	ATC Ile
P		 TAT	CAT	220	TTC		CAG	TCC	CAG	GGT		тсс	CTG	1 GTA	788 GTC	тат	CGG	GGC	AAG	ACC	I GTA	009 TAA									
Val	61y	Tyr	Asp	Asn	Phe	Lys	Gln	Ser	Gln	Gly	Thr	Ser	Leu	Val	Val	Tyr	Arg	61 y	Lys	Thr	Val	•									
			18	19	and the state	182	9		1839		. 1	849		18	59	• • ••••••	186	9	- A 1197	1879		1 מימידינייי	689 GCA	GGGA	18	99 197 1	2002	190 CGAT	9 A		
		GTCC	CGTG	ce c	STACC	GCCC	A GC	CITI	TrgC	GCI	TCAG	AAG	CCGG	1-100	<u></u>			0 AU	CHII	1010			000	339M		•••		700	 a		
		AGTT	19 CCCT	19 GA C	ANGT	192 GCCG	9 T GT	TITC	1939 CTGT	CAG	1 GGAA	.949 .CGT	TTTC	19 CGAG	59 GA 0	CAGO	196 ATTC	C V1	GAAC	CACC TAVA	AGI	TGCA	TGA	ATAC	CCCV	TC /	AGAA	VCCV	c		

Fig. 2. DNA sequences of *T. ferrooxidans rbcL* and *rbcS* and deduced amino acid sequences of the large and small subunits of RuBisCO. The first reading frame (1-1422) corresponds to *rbcL*, and the following one (1453-1809) to *rbcS* (see text for details).

endosymbiont gene products. An even higher degree of similarity may be obtained when comparisons are made on the basis of functional groups of the amino acids (see Table I). It is interesting to note that the highest percentage of similarity both at nucleotide and deduced amino acid levels of rbcL and rbcS genes from T. ferrooxidans is observed with metabolically related bacteria, since C.

vinosum and A. hessleri endosymbionts are also sulfur oxidizing autotrophic bacteria. In contrast, a lower percentage of similarity both at nucleotide and amino acid levels was found for genes and gene products from two purple non-sulfur bacteria, A. eutrophus and R. rubrum (see Table I). It may be recalled that the R. rubrum enzyme has a type II structure [4].

A.	
T.L. C.u	
6.V.	
1	
10	NADETTAX DE D D MK KEME DE VD V I. R D PV A G A V R AC H RAK VDP NNPEO
P. r.	DOSSE U LALK E LIAGGERV CAVIME & YGYVATA HF NVE C TD FTRGV ALV EVDEARELTK.
No. L 4	
T.f.	FYAFIAYPIDLFEEGSVVNVFTSLVGNVFGFKAVRALRLEDVRFPIAYVKTCGGPPHGIQVERDIMNKYGRPLLGCTIKPKLGLSAKNYGRA
C.v.	M N A M
A.b.	I M.N
A.D.	YF L TILI IS IVLFQ LL M
å.e.	FCYV DLS IA LTA II S PIKA M V FA ST I ERLDF A T GR V
R.r.	VA DRNITDGKAMIASFL LTM NO MGD EYAKME FYV E RALFD SVN SALWKVLGRPEVD GLVV TI RP PFAE
	TITLE ******* ******* ****** **
T.f.	CYEGLRGGLDFTKDDENVNSQPFMRWRQRFDFVMEAIQKAEAETGERKGHYLNVTAPTPEQHYKRAEYAKE.IGAPIIMHDYITGGFCANTGLANWCRD
C.v.	
A.h.	VC HGER DEF. TQ
A. n.	VC I Q D L AD B SQ I C E M F .L M FLA T T K
A.e.	V K M I B D L D VN S A V S G M E R F S.L SV V L V. WICIQSMS Q
R.r.	.HAFWL G I N PQGN APL DTIAL AD MRR QD A LFSA I DD FEIIA G VL TF ENASHVALLVD YV GAAAITTARR F
T.f. C.v.	377 IIIII 377 NGMLLHIHRAMEWVLDR.NPHEGIEFRVLTKILRLSGGDHLHSGTVV.GKLEGDREA.TLGWIDIMRDRFIKEDRSRGIFFDQDWASMPGVMPVASGGIEV V A. V A. I. AS. LL ESY L G AFA
A.b.	
8. 0.	
A.C.	DIL G GITT, GAN VS IAW AV AI A. FLIVY HAVE ALLEL I SAGE AND
R.F.	PUNE I G GAVIORYGAN HAF DE DE Y MOET I NOF A LEDDALAMMATTUR LA L'AND DE LEDD - AMA
	473
T.f.	whmpalvtifgd.dsvlofgggtlghpwgnakgaaanrvaleacvearnrgvaiekegkavlteaakhspe.lkiametwkeiksefdtvdkldvahk
C.v.	AC Q DKAS. F I
A.h.	E.HA . E HEL DI IQ . T F
A.n.	E.F PT AQERDLYR GDIR GW.AALDL FEM
λ.e.	GQ EQ IHL . V I Q IQA T M L TE RD LN PEI RDPRAGAAP RARARY GD TFNYTPT TS FVPTASVA
R.r.	LR GFFENL NANVI TA AF ID PVA RSL Q WQWRD PVLDYAREHKELARAF SFPGDADQ YPGWRKALGVEDTRSALPA
в.	100
T.I.	
C.v.	
A.b.	L Nov 3 K F L GVER A L $-C$ $-C$ $-C$ $-C$ $-C$ $-C$ $-C$ $-C$
A.D.	MAT ABAR F F BURN AN 15 MILL TA BIT TA BE FAST A BEAMAYER D'AN ADDIDITION AND AND AND AND AND AND AND AND AND AN
A.O.	KIT Geesessesses a to bey the budy hours above and windshift with the

118 NFKQSQGTSLVVYRGKTV T.f. YA KAEM L P C.v. AM IF PISAKC

¥А A.h. I C TV FI H ... PGRY A.n.

STHTVESVVMSFIVNRPADEPGFRLVRQEEPGRTLRYSIESYAVQAGPK **J.e.**

Fig. 3. Comparison of amino acid sequences of the RbcL (A) and RbcS (B) proteins from T. ferrooxidans (T.f), Chromatium vinosum (C.v). Alvinoconcha hessleri endoxymbiont (A.h), Anacystis nidulans (A.n), Alcaligenes eutrophus (A.e) and Rhodospirillum rubrum (R.r). Only residues that differ from T. ferrooxidans sequence are presented in the one letter code for the amino acids. Residue numbers refer to the T. ferrooxidans sequence. Deletions are indicated by dots. Conserved regions containing amino acid residues participating in catalysis (□) and in assembly of the holoenzyme (***) are indicated.

3.4. Comparison of the deduced RbcL amino acid sequence of T. ferrooxidans with the sequences found in other organisms

Several highly conserved regions became evident by the alignment of amino acid sequences of different bacterial RuBisCO large subunits (Fig. 3). They include amino acid residues which participate both in the active site and in the assembly of the holoenzyme.

The amino acids participating in the active site of the

large subunit of RuBisCO from several bacteria, as reported by Tabita et al. [23], were compared with those from T. ferroxidans (see Fig. 3A). Glu⁵³, a highly conserved residue involved in catalysis [22] defines a conserved region from Ala⁴⁸ to Thr⁶⁴. Lys¹⁷¹ and Lys¹⁷³ belong to a second conserved region starting at Thr¹⁶⁹. The Lys residues have been implicated at the active site of RuBisCO by affinity labeling [22]. The conserved Asp¹⁹¹ associated to a CO₂ binding site and Lys¹⁹⁴ identified as

Table I

Sequence comparisons of the *rbcL* and *rbcS* genes of *T. ferrooxidans* and other prokaryotes. The percent similarity is calculated for amino acids and nucleotides at identical positions in the sequences being compared (aligned as shown in Fig. 3).

		rbcL		rbcS					
		Amino	o acids		Amino acids				
Species	Nucleo- tides	Identical	Related*	Nucleo- tides	Identical	Related			
C. vinosum	82.0	90.0	94.7	72.0	72.0	81.3			
A. hessleri (endosymbion	81.3 t)	87.5	93.0	72.3	71.2	82.2			
A. nidulans	70.8	73.1	86.2	58.6	36.9	55.8			
A. eutrophus R. rubrum	64.6 54.4	55.0 19.7	72.1 29.4	54.0	22.0	31.3			

*Residues belonging to the same functional group are considered related. These groups are: neutral, slightly hydrophobic: P, A, G, S, T; hydrophilic, acid amine: E, D, Q, N; hydrophilic basic; H, R, K; hydrophobic: L, I, V, M; hydrophobic aromatic: F, Y. W; cross-link forming: C.

the site of carbamylation during activation of the enzyme [24], correspond to another conserved region. Another constant region starting at Gly^{281} includes His^{291} , a highly conserved residue also involved in CO₂ binding [22]. The region that includes Lys^{327} was also highly conserved. Affinity-labeling experiments have shown that this residue participates in the active site [25]. Conserved regions starting at positions 100, 214, 235 and 248 corresponded with assembly domains of the large subunit of RuBisCO from several other bacteria [22]. (Fig. 3B).

3.5. Comparison of the deduced RbcS amino acid sequence of T. ferrooxidans and other organisms

Fig. 3B shows the alignment of amino acid sequences of several bacterial small subunits of RuBisCO. With the *rbcS* there is higher divergence than that observed among the large subunits, especially at the amino and carboxy termini. Three highly conserved regions containing amino acid residues related with the oligomer assembly: Glu²⁰, Ser²³, Leu²⁸, Tyr⁶¹, Trp⁶², Pro⁶⁸ and Tyr⁹⁹ [26,27], are shown in Fig. 3B. As expected, the conserved sequence of 16 amino acids that have been proposed to constitute a domain required to facilitate assembly of the holoenzyme from higher plants [15], is absent in *T. ferrooxidans*.

Experiments to study the regulation of the expression of *rbcL* and *rbcS* genes in *T. ferrooxidans* are now in progress.

Acknowledgements: We thank Dr C.C. Allende for criticism of the manuscript. This work was supported by grants from FONDECYT-Chile, by UNDP Proyect CHI/85/002 and by the University of Chile.

REFERENCES

- Murr, L.E., Torma, A.E. and Brierley, J.A. (1978) Metallurgical Applications of Bacterial Leaching and Related Microbiological Phenomena, Academic Press, New York.
- [2] Ingledew, W.J. (1982) Biochim. Biophys. Acta 683, 89-117.
- [3] Holuigue, L., Herrera, L., Phillips, O., Young, M. and Allende, J.E. (1987) Biotech. Appl. Biochem. 9, 497-505.
- [4] Tabita, F.R. and McFadden, B.A. (1974) J. Biol. Chem. 249, 3459-3464.
- [5] Gibson, J.L. and Tabita, F.R. (1977) J. Biol. Chem. 252, 943-949.
- [6] Gibson, J.L. and Tabita, F.R. (1977) J. Bacteriol. 132, 818-823.
- [7] Shinozaki, K., Yamada, C., Takahata, N. and Sugiura, M. (1983) Proc. Natl. Acad. Sci. USA 80, 4050-4054.
- [8] Viale, A.M., Kobayashi, H. and Akazawa, T. (1989) J. Bacteriol. 171, 2391-2400.
- [9] Curtis, S.E. and Haselkorn, R. (1983) Proc. Natl. Acad. Sci. USA 80, 1835-1839.
- [10] Nargang, F., McIntosh, L. and Somerville, C. (1984) Mol. Gen. Genet. 193, 220-224.
- [11] Zurawski, G., Perrot, B., Bottomly, W. and Whitfeld, P. (1981) Nucleic Acids Res. 9, 3235-3270.
- [12] Stein, J.L., Haygood, M. and Felbeck, H. (1990) Proc. Natl. Acad. Sci. USA 87, 8850-8854.
- [13] Andersen, K. and Caton, J. (1987) J. Bacteriol. 169, 4547-4558.
- [14] Shinozaki, K. and Sugiura, M. (1983) Nucleic Acids Res. 11, 6957-6964.
- [15] Wasmann, C.C., Ramage, R.T., Bohnert, H.J. and Ostrem, J.A. (1989) Proc. Natl. Acad. Sci. USA 86, 1198-1202.
- [16] Kusano, T., Sugawara, K., Inoue, C. and Suzuki, N. (1991) Curr. Microbiol. 22, 35-41.
- [17] Mackintosh, M.E. (1978) J. Gen. Microbiol. 105, 215-218.
- [18] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- [19] Tabita, F.R. and Small, C.L. (1985) Proc. Natl. Acad. Sci. USA 82, 6100-6103.
- [20] Hanahan, D. and Mcselson, M. (1983) Methods Enzymol. 100, 333-342.
- [21] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
- [22] Tabita, F.R. (1988) Microbiol. Rev. 52, 155-189.
- [23] Hwang, S. and Tabita, F.R. (1991) J. Biol. Chem. 266, 6271-6279.
- [24] Miziorko, H.M. and Lorimer, G.H. (1983) Annu. Rev. Biochem. 52, 507-535.
- [25] Stringer, C.D. and Hartman, F.C. (1978) Biochem. Biophys. Res. Commun. 80, 1043-1048.
- [26] Fitchen, J.H., Knight, S., Andersson, I., Branden, C. and McIntosh, L. (1990) Proc. Natl. Acad. Sci. USA 87, 5768-5772.
- [27] Lee, B., Berka, R.M. and Tabita, F.R. (1991) J. Biol. Chem. 266, 7417-7422.