Primary structure of an archaebacterial transducer, a methyl-accepting protein associated with sensory rhodopsin I

(phototaxis/chemotaxis/signal transduction/halobacteria)

VIRGINIA J. YAO AND JOHN L. SPUDICH

Department of Anatomy and Structural Biology, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461; and Department of Microbiology and Molecular Genetics, University of Texas Medical School Health Science Center, Houston, TX 77030

Communicated by Walther Stoeckenius, September 17, 1992

ABSTRACT A methylated membrane protein of 97 kDa was suggested on the basis of mutant analysis to transduce signals from the phototaxis receptor sensory rhodopsin I to the flagellar motor in Halobacterium halobium. Here we report isolation of the proposed transducer protein, cloning of its gene based on partial protein sequences, the complete gene sequence, and analysis of the encoded primary structure. The 1611-basepair gene termination codon overlaps the initiator ATG of the sopI gene, which encodes the sensory rhodopsin I apoprotein. The predicted size of 57 kDa for the methylated protein indicates an aberrant electrophoretic migration on SDS/polyacrylamide gels, as occurs with other acidic halophilic proteins. Putative promotor elements are located in an A+T-rich region upstream of the gene. Comparison of the translated nucleotide sequence with N-terminal sequence of the purified protein shows the protein is synthesized without a processed leader peptide and the N-terminal methionine is removed in the mature protein. The deduced protein sequence predicts two transmembrane helices near the N terminal that would anchor the protein to the membrane. Beyond this hydrophobic region of 46 residues, the remainder of the protein (536-amino acid residues total) is hydrophilic. The C-terminal 270 residues contain a region homologous to the signaling domains of eubacterial transducers (e.g., Escherichia coli Tsr protein), flanked by two regions homologous to the methylation domains of the transducer family. The protein differs from E. coli Tsr in that it does not have an extramembranous-receptor binding domain but instead has a more extended cytoplasmic region. Coexpression of the methylaccepting protein gene (designated htrl) and sopl restores sensory rhodopsin I phototaxis to a mutant (Pho81) that contains a deletion in the htrl/sopl region. These results extend the eubacterial transducer family to the archaebacteria and substantiate the proposal that the methylated membrane protein functions as a signal-transducing relay between sensory rhodopsin I and cytoplasmic sensory-pathway components.

Sensory rhodopsin I (SR-I) is a retinal-containing intrinsic membrane protein that functions as a phototaxis receptor in the archaebacterium *Halobacterium halobium* (also known as *Halobacterium salinarium*) (1). SR-I controls swimming behavior of the cells by modulating the frequency of reorientation ("reversals") of their swimming direction (for reviews, see refs. 2 and 3). Orange light generates attractant signals that suppress reversals, whereas near-UV light generates repellent signals that induce reversals. Several early events in the signaling process have been elucidated. Photon absorption by SR-I (λ_{max} , 587 nm) causes isomerization around the C13—C14 double bond of the retinal chromophore (4), which is essential for receptor activation (5). The photoisomerization energy is transferred to the protein in a process requiring steric interaction between the retinal C13 methyl group and protein residues (6). These events produce in <1 msec a spectrally distinct attractant signaling state (λ_{max} , 373 nm) of the protein that decays thermally within seconds to the original unstimulated-receptor conformation (7, 8). The repellent signal is generated by absorption of a second photon by the receptor when it is in its attractant signaling conformation (9).

What is the nature of the components responsible for transmission of these signals to the flagellar motor? The gene-derived primary structure of SR-I reveals a seventransmembrane helix protein largely embedded in the membrane that has minimal cytoplasmic loops, similar in its structure to the related proton pump bacteriorhodopsin (10). Unlike bacteriorhodopsin, SR-I is not an electrogenic ion pump (1, 11), and photoreceptor signals are not propagated as electrical impulses along the membrane (12). Fumarate (13), cyclic nucleotides and Ca²⁺ (14), and a G protein (15) have been suggested to be involved in postreceptor signaling. However, no postreceptor component capable of sensing SR-I conformations and transducing this information to modulate the motor has been demonstrated in these studies.

Phototaxis mutant analysis identified a methylated membrane protein with an apparent molecular mass of 97 kDa, expression of which tightly correlated with the SR-I protein of 25 kDa (16). The methylation linkage (reversible carboxylmethyl esterification; refs. 16 and 17) of the 97-kDa protein is of the same type as that found in the chemotaxis signal generators ("transducers") of eubacteria (e.g., Escherichia coli), chemoreceptor proteins that transmit signals from the membrane to a cytoplasmic sensory pathway (18-20). SR-I attractant and repellent signals modulate methyl group turnover in vivo, as do those from eubacterial chemotaxis transducers (21). Based on these findings the 97-kDa methylaccepting protein was postulated to be the postreceptor transducer for SR-I signals (21). Analysis of taxis mutants and revertants (22, 23) and antigenic crossreactivity of the 97-kDa protein with eubacterial transducers (24) further strengthened the analogy to the chemotaxis system.

Here we report the cloning and sequencing of the gene encoding the proposed transducer protein.* The protein primary structure and its sequence similarities to eubacterial transducers substantiate the proposal that the 97-kDa protein acts as a signal-transducing relay between SR-I and cytoplasmic sensory pathway components. In this paper we refer to the gene as *htrI* and the protein as HtrI, in reference to its function as a halobacterial transducer for SR-I. In earlier papers we referred to HtrI as methyl-accepting phototaxis protein I (MPP-I).

MATERIALS AND METHODS

Bacterial Strains. *H. halobium* Flx5R (HtrI⁺SR-I⁺), which produces 4-fold greater amounts of HtrI than does a wild-type

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: SR-I, sensory rhodopsin I.

^{*}The sequence reported in this paper has been deposited in the GenBank data base (accession no. L05603).

(25) strain was used for HtrI isolation and preparation of genomic clones. Pho81 (22) [HtrI⁻SR-I⁻, resulting from an insertion of the transposable element ISH2 into position 1775 of *sopI* and an upstream deletion including *htrI* (E. N. Spudich and J.L.S., unpublished work)] was used as a negative control. *E. coli* DH5 α (*supE44* Δ *lacU169* (ϕ 80 *lacZ* Δ *M15*) *hsdR17 recA1 endA1 gyrA96 thi-1 relA1*) was used as the host strain for cloning. *H. halobium* DNA was cloned into the plasmid vector pBluescript II KS(+) (Stratagene).

Molecular Biological Procedures. Standard protocols were used (26). Enzymes were purchased from GIBCO/BRL, New England Biolabs, Boehringer Mannheim, Promega, and United States Biochemical. Exonuclease digestions were done by using the Erase-a-Base kit (Promega) following the manufacturer's protocol.

Protein Isolation. H. halobium strains Flx5R and Pho81 were grown to stationary phase at 39°C in complex medium, and sonicated membranes were isolated (16) and dialyzed against water at 4°C for 1.5-2 hr. The suspension was solubilized, and proteins were separated electrophoretically on an SDS/low-bisacrylamide/polyacrylamide gel, as described (16), except the gels were polymerized for 48 hr at 4°C, preelectrophoresed in 10 mM glutathione at 5 mA constant current for 2 hr, and electrophoresed overnight at 6 mA. The gels were stained with Coomassie brilliant blue R-250, and a small portion of the gel was subjected to autofluorography with EN³HANCE (DuPont/NEN). Comparison of the autofluorogram with the stained gel and absence of the band in Pho81 identified the 97-kDa band. Slices containing protein from three to four gels were pooled, macerated with a spatula, and soaked in 0.4 mM NaHCO₃/2% SDS for 20 min at room temperature. Protein was electroeluted from the gel matrix for 16 hr at 50 V in 50 mM NH₄HCO₃/0.1% SDS, dialyzed for 8-12 hr against 1 liter of 5 mM NH₄HCO₃/0.01% SDS (twice), lyophilized, and resuspended in 50 μ l of water. Two and one-half microliters of the resuspension was applied to a 10% acrylamide/0.27% bisacrylamide mini-gel (Hoefer) to confirm its purity and identity by autofluorography, by migration position relative to molecular weight standards (Bio-Rad), and by comparison to the [3H]methylated Pho81 and Flx5R membrane proteins. The remaining 45 μ l of protein suspension was precipitated with 0.475 ml of 1 mM HCl/acetone solution at -20°C for 2 hr to remove residual SDS (27). The protein pellet was dried in a Speed-Vac evaporator and stored at -20° C. Dried protein from 24 preparative gels was solubilized and pooled in 50 µl of 31 mM Tris·HCl, pH 6.8/2% SDS/5% (vol/vol) glycerol/0.001% bromophenol blue/5% (vol/vol) 2-mercaptoethanol, and the resuspension was gently sonicated 30 sec twice with ice-bath cooling for 30 sec between sonication periods (Branson model 1200 sonicator). The solubilized protein was applied to a preelectrophoresed SDS/minipolyacrylamide gel. Separate protein samples were electroblotted onto poly(vinylidene difluoride) (Bio-Rad) and nitrocellulose (Schleicher & Schuell) membranes at 2.5 mA/cm² with a semi-dry electroblotter (Integrated Separation Systems, Hyde Park, MA or Bio-Rad). Membranes were stained for 1.5 min in 0.2% Ponceau S in 1% acetic acid, destained for 1 min in 1% acetic acid, and rinsed with milliQ (Millipore, Bedford, MA) water for 1.5 min. The stained protein was excised from the blot, rinsed extensively with water in a 1.5-ml microfuge tube, frozen at -20° C, and sent on dry ice to the Harvard Medical School Microsequencing Facility (Boston)

Peptide Sequencing. Ten percent of the nitrocellulose and of the poly(vinylidene difluoride)-immobilized protein was quantitated by amino acid analysis. The nitrocellulose-blotted sample was digested with trypsin, and peptides were separated by reverse-phase HPLC by a Vydac $C_{18} 2.1 \times 15$

cm column. Sequence was obtained from two peaks as well as from the N terminus of the poly(vinylidene difluoride)blotted protein with an Applied Biosystems 477A protein sequencer with a 120A online phenylthiohydantoin-amino acid analyzer.

Deoxyoligonucleotides and Southern Hybridization. Three fully degenerate (64-, 256-, and 512-fold) deoxyoligonucleotide mixtures were designed and synthesized (Genosys, The Woodlands, TX) based on the 28-amino acid tryptic peptide sequence (Fig. 1). The deoxyoligonucleotides were endlabeled with T4 polynucleotide kinase (GIBCO/BRL) and $[\gamma^{-32}P]$ dATP (6000 Ci/mmol, DuPont/NEN; 1 Ci = 37 GBq). Restricted Flx5R genomic DNA was blotted onto Hybond-N (Amersham) by either the Southern blot capillary method or the Pharmacia VacuGene system. Southern blots were prehybridized in sealed plastic bags with 20 ml of $6 \times SSC/0.05$ M NaPO₄, pH 6.8/50 mM EDTA, 5× Denhardt's solution/ 200 μ g of denatured calf thymus DNA at 37°C for at least 4 hr. The prehybridization buffer was discarded, and 10 ml of fresh buffer containing labeled probe was added. After a 66-hr incubation at 37°C, blots were washed three times for 15 min with 150 ml of 3× SSC/0.1% SDS. Hybridization stringencies were increased by decreasing ionic strength $(3 \times -0.1 \times SSC)$ and increased wash-solution temperature (37°C-65°C) until single hybridization signals were obtained with the 64-fold degenerate probe. The 11-kilobase-pair (kbp) EcoRI fragment (Fig. 2) was confirmed by its high-stringency hybridization to each of the other two probes. Blots were exposed to either Kodak X-Omat XAR film or Hyperfilm (Amersham).

Plasmid Preparation. Overnight E. coli cultures (1.5 ml) were grown in LB broth containing 100 mg of ampicillin per liter. Plasmids were prepared by using the Magic Mini Plasmid Prep kit (Promega). DNA was eluted from the cartridge with sterile milliQ water at 68°C, and the final volume was adjusted to 50 μ l with additional milliQ water. Plasmids were denatured at room temperature for 5 min by adding 5 μ l of a 2 M NaOH/2 mM EDTA solution. DNA was precipitated by adding 7.5 µl of 3 M NaOAc (pH 5.2), 7.0 µl of sterile milliQ water, and 200 μ l of -20° C absolute ethanol. The preparation was incubated at -20° C for at least 20 min and centrifuged at 4°C for 15 min at 14,000 \times g, the supernatant was aspirated, and the pellet was rinsed with 70% (vol/vol) ethanol/30% (vol/vol) Tris/EDTA, pH 8.0. The final pellet was dried in the Speed-Vac evaporator, resuspended in 25 μ l of Tris/EDTA, pH 8.0; 5 μ l was then used for each hybridizing reaction.

Sequencing. The labeling reactions were prepared by using a modified protocol of the Sequenase version 2.0 kit (United States Biochemical). Hybridizing reactions contained 5 μ l of denatured plasmid, 2 μ l of Sequenase reaction buffer, 2 μ l of primer (20 ng), and 1 μ l of dimethyl sulfoxide (Fisher Scientific), incubated at 65°C for 10 min and cooled immediately in an ice/slush bath. The hybridized template $(2.5 \mu l)$ was added to 4 wells of a round-bottomed microtiter plate (Falcon), and the plate was incubated on a 45°C heating block. One microliter of labeling mix $[1 \mu]$ of 0.1 M dithiothreitol, 1.6 μ l of sterile water, 0.4 μ l of dimethyl sulfoxide, 0.4 μ l of dGTP or dITP sequence-labeling mix, 0.5 μ l of deoxyadenosine 5'- $[\alpha-[^{35}S]$ thio]triphosphate (1000–1500 Ci/mmol, DuPont/ NEN), 0.25 μ l of Sequenase version 2.0] was added to the side wall of each well and mixed in a Mistral 2000 centrifuge (Leicestershire, England). The dITP mix included 0.125 μ l of pyrophosphatase and 1.475 µl of sterile water. After centrifugation, the microtiter plate was returned to the 45°C block for 8 min. Two microliters of each Sequenase dideoxynucleotide chain-termination mix was added to each well, centrifuged to mix, and the plate was returned to the 45°C block for 10 min. Three microliters of stop solution was added to each well and mixed; the quenched reactions were stored at

Biochemistry: Yao and Spudich

-283 -164 -44		ATC	CG AAA CTG	TCC ACI	GT TG CC	TG	GGA TAT		AAC ATZ		GCO ATI	GGT ACC GGA	CG AC	TAC AAG TGA	AT AA TA	GTC GAG AC1	GC GAG	AGG TGT TGG	GT TGC	rgg CCG ATG	GGG GCT	GAC FGC AAT	CG TC CG	TCC GGC CA1	CAG CGC CGG	GA	CAG ACC GCG	TTC AGC	GCA	GGT ATA GAT	TAC	CCA IGT GGC	GAC CAA GTT	GTA CTC	AT TC	CAC CGT	AG1 AGC GC1	TTG CAG	SCG SGT	CTG TTA ATA		GCC	GTC GTT ACC	CGT TTC GCC	TTC ATC	CCC	TCC TGA CCTG
:	L																		ľ	1 _	Т	I	A	ţ.	1	A	R	R	R	Y	· (3	V	к	L	G	L	G	Y	I	P	<u>،</u> ۱	r	A	G	L	L
70 20	5 G 5 V	TCC	GG(GTC V	:GG G	CG: V	rc <i>a</i> T	cc	AC <i>I</i> T	N N	CGI D	ACG V	TCO	200	TC S	GAC T	CA: I	ГТG V	TCC	GCC A	GGI G	AAT I	CG A	ccc G	GG ;	TT <i>i</i> L	ACT L	CAC T	GC: L	rcg G	GAT	rcg. S	ATC I	AAC N	GCC A	GC A	CGA E	GAC T	CCG V	TCG A	CCA	GCI	ATC I	AAG K	GAG E	SAT(I	CGCC A
190 60	i C i A	CAC	AG	ACC F	GA E	GCC R	GCG V	TC	GCC A	CAA N	CGC G	GGA N	ACO	CTC	GA. E	ACA Q	GGI E	AGG V	TG i 1	CC C	TCC S	CAC T	CC(R	GAA I	CG	GAC D	CGA E	GTI F	CGC G	GAT S	CGC I	стто , 1	SCC	GAC D	TCC S	I I	CGA E	ACZ Q	AGA' M	IGC R	GCC Q	AG1	rcg S	CTT(L	CG1 R	GGC G	CCGC R
316	c	TGA	AC	GAG	AT	GGI	AGC	GC	ACC	cG	cGC	cG	ACO	стс	GAI	AGA	GAC	CGC	AGG	cc	GAC	GC	cGi	AGA	CG	GCC	cG	CGA	GGI	AGG	cco	AG	CAA	GCC	AAC	CAC	GGA	AGO	:cci	AGG	ccg	cGG	AG	CGC	GAG	GCC	ccc
106	о <u>г</u>	N		5	м	E	R		Ľ	ĸ	A	D	1		E	E	T	Q	P	1	E	A	E	T		A	ĸ	E	E	A	Ŀ		2	A	ĸ	Q	E	A	Q	A	A	. <u>r</u>		× 1	5	A	ĸ
436 146	G E	AGC	TGC	GCA A	GC(A	CAC T	CT Y	AC	CAG 2	GA D	CAC T	CG A	CC7	lag (CG(R	Y Y	GGC	E E	AAA T		ATC M	GA E	GG(A	CCG A		GCC A	T	GGG G	GG# D	L L	TCA T	CCC C		CGG R	GTC V	GAC D	CGT V	GGA D	CAC T	CGI D	ACC H	ACG E	SAGO	GCG7	ATG 1	GAG E	ACG T
556 186	G V	TCG G	GC# 1	CC.	GC(A	GTI F	CA N	ACO	CAG	AT M	GAT M	GG. D	ACC E	SAC	CTC L	CA Q	GGC A	GA T	CGG V	TC	CGA R	AC T	GG1 V	rca T	ر22. 1	ACG F	GT V	CGC A	CGA D	ACGI E	AGA I	TCC	AA(GCC. A	AAA K	ACC T	GA E	GCG R	GA1 M	IGTO S	CCG E	AGA T	CC1	4GC0 5 /	GCC A	GAC D	ATC I
676	G	AGG	cc,	GC	GCC	CGG	TG.	AC/	cc	GT	CGA	AG	CGG	TG	rce	AA	GAI	CG	AGT	CG	CAG	GC	CAP	ACG	ACC	CAG	CGG	CAC	GGA	GC	rcg	АСІ	CGG	GCG	GCC	GAC	GA	CGT	ccł	\GC1	AGG	TCA	GTC	GCG1	rcc	GCC	GAG
226	E	A	S		A	G	D	3		v	Е	A	v		s	ĸ	I	Е	s	(5	A	N	D	ç	2	R	т	Е	L	D	s	1	A 1	A	D	D	v	Q	Q	v	s	1	1 5	5.	A	E
796 266	Gž E	AGA I	TCG A		GCC A	CAC T	CA' I	rco E	AC	GA0 D	L L	GGC A	CGT S		CGC R	AG S	CGA E	GGI D	ACG V	TAC 1	GCG A	AC(T	CGC A	CA S	GCC E	SAC	GCC A	GGC A	CCG R	CGI D	ACT S	сст s	CG/	AG	rcc 5	GCG A	CT L	CGA D	CGA E	IGAT M	rgt(S	CCA S	GC I	ITCG	GAG.	ACC T	GAG E
916	G	rcg.	ACG	AC	sco	GT	CG	SCC	AG	GTO	GA		AGC	TGO	CGC	GAG	CA	GGI	cĢ	cco	GAG	ATO	CAC	cGi	ACA	TC	GTO	GA	CGT	GAI	CA	ccc	ACA	TCO	GC	GAG	CAG	GAC	GAA	CAT	GC	rgg	cGC	TGA	AC	GCG	тсс
306	v	D	D		4	v	G	ų		v	е —	Q			< 		Q		- A	E	-	1	т 		4		v 		·		т	ں _			; 	Ľ.	Q	т	N	м	г	A	1	, N		A	5
1036 346	A] I	CG. E	AGG A		GCC A	CG R	CG(A	CCG G	GC	GGC G	CAA N	CGC A	CG: D	ACC	GGG GGG	GA1 D	rGG G	CTI F	CA S	GCC N	STC /	GT(V	CGC A	CGI D	ACG E	AG	GTC V	CAAC K	GGA D	L L	CG A	CGG E	AGG E	1AA	ICG	CAG 2	GA(D	CCG R	CGC A	CAA N	CGI E	AGA' I	TCG A	CCG	CAG	GTG V	GTC V
1156	G/	GA	٩AG	TC	ACC	:GC	cci	AGA	cco	GAG	GA	CGI	CA	cco	CG	тсо	CAT	cc7	GC	AG2	ACC(CGł	AC	cco	GGG	TG	GAG	TCC	GG	CAG	CG	-	CCG	TCC	SAA'	rcc	ACC	CT	GCG	GGA	CAT	100	GCA	CGA	тсо	GCG	GAC
386	E	ĸ	v	5	r	A	Q	Т	1	E	D	v	т	7	1	s	I	Q	Q	1		R	Т	R	v	' 1	E	s	G	s	E	Т	v		: :	5	т	L	R	D	I	R	т	I	2	A 1	C
1276 426	тс s	CA:	rcg A	CTC E	GAG	GT(V	сто s	GA N	AC:	rcg 5	I I	CGA D	CG. E	AGA I	тс	CAG	CG R	CAC T	GAG	CCI S		GAG E	CA Q	GGC A	CG E	AA	ACC F	GTC V	Q	AAG S	CAC T	CCG A	CCA T	רכם פ	CTO	stg /	GAG E	R R	rgt V	CGC A	GGGG	AC: L	rga S	GCG. D	ACC	GAC/	ACG r
1396 466	AC T	GGC A	GT L	TGO P	cG	AG(S	CGP D	CG A	CCC E	GAG C	TCC S	CGC A	GG V	IGA I	TC	GGG G	CA Q	ACG R	CG/ E	AGA S	GC(GCA A	GA E	GG <i>P</i> E	I I	TCO	GCG A	GCG A	TC S	CCT L	CG/ E	AAC Q	AGT F	тсс	AG		ACG F	GCC A	CGT V	CGA E	GCA Q	GCI	rgc Q	AGT S	CCC F		GTG /
1516	GC	GTC	GT:	ГСА Т	cc	GT(GC	CA T	CCG	GAG	GAC	CAG	CGI E	AGA T	CT	GCG	GGG	GGG	TAC	SCG	TGO	GAG	CA	GCC	AG:	TCI	ATG	CGI R	GCO	CGG	CGC	TG	ACG	GGG	GTO	GT	GCG	TGZ	TG	GAC	GCC	GTC	GC.	AAC:	CGC	CTA	
1600			•		~~	•		-	-					•				•					*	•		•									~~~									•		•	
1636	TC	GGC	GGGG	A	V	TCC I	acg A	L	I	V	TCC C	GT	V	A	F	V	101	rgg V	L	L	Y	R	GG	rcg S	L	D	G	SCI	ECC	, CG	H	Q	S	A	L	A	P	I	rgo , i	A	ATC	I	P	V	F	A	G
1756 50	GC	CTG L	TCC S	Y Y	CG	TGO	GG	AT(M	GGC A	GT. Y	ACG	SAC.	ATC I	GG G	AA(T	CGG V	TG/]	ATC	GTA V	AA N	CGC G	GA N	ACO	CAG 2	AT(I	CG1 V	CG G	GGC L	TGC F	266 }	TAC Y	ATC I	GA D	CTG W	GCI L	V V	rga T	CGA I	CGG	CGi	ATC I	CTC L	CGTO V	G G	GTA Y	CGI V	CG G
1876 90	GC	TAC Y	GCC A	GC A	GGG G	GGG A	CG	тс¢ s	CCG R	TC R	GCA S	GC	ATC I	TA:	CGC G	STG V	TG# M	TG I	GTG V	GC A	GGA D	CG A	CGC I	стс	ATC M	GAT I	CG A	cgg V	TGG	GC	GCG A	GGG G	GC(A	GT V	GGI V	GAC T	TG: D	ACG G	GC /		CTC	AAG K	TGC W	GCC A	GCT L	GTI F	CG G
1996 130	GC	GTG V	тсс s	STC S	GAT	ГСТ F	тс	CAC H	CT L	GT(S	CGC L	TG	TTC F	GC A	GT/ Y	L CC	IGI Y	AC	STG /	AT(I	CTT F	TC P	CGC F	266 }	GTC V	CGT V	GCO	CCG. D	ACG V	TGC	ccc	GAG E	CAC Q	GAT I	CGG G	GCI L	GT F	TCA N		TGC:	CTG.	AAA K	AAC N	CAC	I I	ccc G	GC L
2116 170	TG(]	CTG	TGG W	TT L	GGC A	CGT Y			CT L	GG: V	IGT W	GGG	CTG L	TTO F	CGG G	CC(P	CGG A		GC GC	AT(I	cGG G	GGI E	AGG A	CA	ACG T	GC A	TG(A	CCG G	GCG V	TCC	SCG	стс ь	ACC T	TA Y	CGT V	GTI F	CT	rgg D	ACG V	;TGC / I	TC	GCG A	AAG K	GTC V	CC P	GTA Y	CG V
2236 210	TGI	TAT (ITC F	TT(F	CT# Y	ACG A	CG	CGG R	CG R	TCC R	SCG V	TG1 F	rtc	ATC M	GCA H	CTC S	CGG E	AGI	CG	CCC P	GCC P	GG(A	CTC P		GAG E	CA Q	GGC A	GA(T	ccg V	TCC	GAG	GCG A	ACC T	GC(A	GGC A	GGA D	СТ(*	GAC	GGG	ccc	STCO	ccc	GGI	GGT	GG	IGT	GA

2356 CCCCGGCTCCGTCCAGCAGTCTGAGAGCCGCAACAAGGTTCAAGACACTGGTGGGCTTCTCTTCAGGTGAAGGTCCATGGATATC

FIG. 1. Nucleotide and encoded protein sequences of the *htrI/sopI* region. The *htrI* gene spans nucleotides 1–1611, and the *sopI* gene spans nucleotides 1611–2331. Underlined amino acid residues were determined by tryptic peptide and N-terminal sequencing. The three half-arrows drawn above the DNA sequence denote regions of the peptide sequence to which fully degenerate deoxyoligonucleotide mixtures complementary to the noncoding strand were synthesized for cloning. Putative promoter sequences are boxed.

 -20° C. Sequencing reactions were loaded onto a preelectrophoresed [0.5× TBE (1× TBE is 90 mM Tris/64.6 mM boric acid/2.5 mM EDTA, pH 7.3), 30 min, 50 W] 5% polyacrylamide gel containing 0.5× TBE and 7.5 M urea and run at 50 W constant power.

RESULTS

Gene Sequence and Mapping. The *htrl* gene was identified on an 11-kbp *Eco*RI fragment of Flx5R genomic DNA by high-stringency Southern hybridization with each of the three degenerate 5'- $[^{32}P]$ phosphorylated oligonucleotide mixtures based on an internal tryptic peptide sequence (Fig. 1). The gene was sequenced in both directions by generating overlapping subclones derived from restriction digests and exonuclease III nested deletions (Fig. 2). The gene-derived primary structure confirmed the two internal tryptic peptide sequences and the N-terminal sequence (Fig. 1).

The termination codon of the 1611-bp htrI overlaps the first base of *sopI*. Two putative promoter elements exhibiting close sequence identity to promoter sequences in other halobacterial genes (28) are located in an A+T-rich region upstream of htrI (Fig. 1).

The Encoded Protein. The *htrl* gene encodes a protein of 536 amino acids with a calculated molecular mass of 56,675 Da and a pI of 3.9. Comparison of the translated nucleotide sequence with N-terminal sequence of the purified protein shows the protein is synthesized without a processed leader peptide and the N-terminal methionine is removed in the mature HtrI. A cluster of four basic residues is present near the N terminal, and a hydropathy plot of the translated gene indicates two hydrophobic segments immediately after these



residues (Fig. 3A). These putative transmembrane regions are followed by a highly hydrophilic and highly acidic structure extending to the C terminus (Fig. 3 A and B).

Comparison of HtrI to E. coli Tsr. A search of the protein sequences in the GenBank and European Molecular Biology Laboratory data bases identified the highest sequence similarity of HtrI to eubacterial chemotaxis signal transducers, [e.g., E. coli Tsr, Tar, Trg, Tap, and related proteins in other eubacteria (18, 20)]. HtrI exhibits 64% identity and 87% similarity (Genetics Computer Group package version 7.1) to a 47-amino acid region (HtrI residues 325-371) within the signaling domain (29) of Tsr (Fig. 4). The flanking methylation domains in Tsr show a lesser degree of sequence identity to HtrI; however, the five doublet carboxylmethylation sites identified in Tsr (30) correspond to glutamate and aspartate residues in HtrI (Fig. 4), potential carboxylmethylation sites. Further, four of these potential sites in HtrI occur in doublet acidic residues, which is characteristic of eubacterial carboxylmethylation sites (30).

DISCUSSION

In the htrI/sopI region, the only putative promoter sequences with strong identity to the *H. halobium* consensus promoter sequence (28) reside upstream of htrI (Fig. 1). Two putative promoter elements with weaker sequence identity to an archaebacterial consensus promoter sequence were identified within 33 bp upstream of sopI (10). To study expression of htrI and sopI in *H. halobium*, we used a selectable vector





FIG. 2. Restriction map of htrl. Key restriction sites on the initially isolated 11-kbp EcoRI fragment are indicated. A 2.7-kbp BamHI/EcoRV fragment contains the htrl/sopl region. The htrl gene was mapped by restriction analysis. B, BamHI; Bs, BstXI; E, EcoRI; Ev, EcoRV; P, Pst I; S, Sal I; Sm, Sma I; Ss, Sst II; and X, Xma III.

(M. P. Krebs, R. Mollaaghababa, and H. G. Khorana, personal communication) derived from a described plasmid (31) combined with the mevinolin-resistance gene (32) to transform strain Pho81. SR-I expression was not detected from the *Bcl I-Sal I* fragment (from 986 to 34 bp downstream of bp 2437, Fig. 1), which contains the suggested *sopI* promoters. However, expression of both HtrI and SR-I was obtained when this fragment was extended upstream to the *Bam*HI site (-283) to include *htrI* and its putative promoters (Fig. 1), suggesting *htrI* and *sopI* are cotranscribed. Pho81 transformants containing the *htrI/sopI* region exhibited normal SR-I photochemical activity, as assessed by flash photolysis and complete restoration of SR-I-mediated phototaxis responses *in vivo* (unpublished work).

Several H. halobium genes exhibit at -7 to -12 a sequence complementary to the 3' end of the H. halobium 16S RNA (ribosome-binding site) (33). Such a properly positioned ribosome-binding site is not evident upstream of htrI, although a possible site occurs at -24 to -20. The importance of ribosome-binding sites in H. halobium is not clear because there are several H. halobium genes that appear to lack them (33).

The apparent relative molecular mass of 97,000 Da for HtrI on SDS/polyacrylamide gels is higher than the 56,675 Da

Htrl	ADDVQQVSASAEEIAATIDDLASRSEDV.ATASDAARDSSKSALD	297
Tsr	GASEIATGNNDLSSRTEQQAASLEETAASMEQLTATVKQNAENARQASHL	
Htrl	EMSSIETEVDDAVGQVEQLRDQVAEITDIVDVITDIGEQTNMLA	341
Tsr	ALSASETAQRGGKVVDNVVQTMRDISTSSQKIADIISVIDGIAFQTNILA	
Htrl	LNASIEAARAGGNADGDGFSVVADEVKDLAEETQDRANEIAAVVEKVTAQ	391
Tsr	LNAAVEAARAGEQGRGFAVVAGEVRNLAQRSAQAAREIKSLI	
Htrl Tsr	TEDVTASIQQTRTRVESGSETVESTLRDIRTIADSIAEVSNSIDEIQRTT : :: :: : : : : : : .EDSVGKVDVGSTLVESAGETMAEIVSAVTRVTDIMGEIASASDEQSRGI	441
Htrl	SEQAETVQSTATSVERVAGLSDDTTALASDAESAVIGQRES : : : : :: : :	482
Tsr	DQVGLAVAEMDRVTQQNAALVEESAAAAAALEEQASRLTEAVAVFRIQQQ * *	
Htrl	AEEIAASLEQFONTA 497	
Tsr	QRETSAVVKTVTPAAP *	

FIG. 4. Primary-sequence comparison of HtrI residues 254-497 (Fig. 1) to the methylation and signaling domains of E. coli Tsr (29). Stars identify carboxylmethylated glutamyl residues in mature Tsr (30).



FIG. 5. Schematic model of HtrI. The N-terminal residue is indicated as the second of the 536 encoded residues in accord with N-terminal sequence data (Fig. 1). The 56th residue corresponds to the end of the two hydrophobic segments, and the 265th residue corresponds to the beginning of the region homologous to Tsr.

calculated for the mature protein from the translated gene sequence. Anomalously slow migration in an SDS/ polyacrylamide gel is expected from the highly acidic nature of the protein and has been observed in the acidic H. halobium ATPase α and β subunits (34) and cell-surface glycoprotein (35). This effect may be sufficient to explain the slow migration; however, we have not ruled out that the molecular mass is influenced by posttranslational modification, dimerization, or binding of an additional component resistant to SDS and reducing agents.

A common motif found in diverse eubacterial signal transducers is a highly conserved signaling domain in the C-terminal portion flanked by two methylation regions that function in adaptive attenuation of the signal (18, 20). Existence of this motif in HtrI extends the family of eubacterial signal transducers to the archaebacteria, indicating an early origin for this type of signal transducer. In eubacterial chemotaxis, the signaling domain controls a two-component regulatory system that is a member of a widespread family of cytoplasmic phosphotransfer systems (19, 20). Such a cytoplasmic transduction system has not been demonstrated in archaebacteria, but the existence of a similar signaling domain on HtrI suggests that this system may occur. In this regard, phenotypes of phototaxis and chemotaxis mutants of H. halobium exhibit similarities to first- and second-component mutants in E. coli chemotaxis (22, 23).

A model for HtrI can be constructed based on sequence similarity to eubacterial transducers, hydropathy-plot analysis, and the requirement of a cytoplasmic location of the methylation and signaling domains (Fig. 5). The distribution of charges surrounding the putative transmembrane segments is as expected from the view that charged residues act as a stop transfer signal (36) and from the presence of a transmembrane electrical potential in H. halobium that stabilizes positive and negative charges at the intracellular and extracellular membrane surfaces, respectively. HtrI differs from E. coli Tsr in that HtrI does not have an extracellularreceptor binding domain but, instead, has a more extended cytoplasmic region. This structural distinction fits the different functions of the two transducers. Tsr is a chemoreceptor interacting with extracellular effectors, whereas HtrI must sense the conformational states of SR-I.

Physical proximity of HtrI to the retinal-binding site of SR-I has been suggested by a possible migration of [3H]retinal

from SR-I to HtrI during a borohydride-reduction procedure (16). An attractive possibility is that the cytoplasmic loop (residues 56-265) is positioned at the cytoplasmic surface of the SR-I protein. Physical proximity would concur with preliminary results from SR-I expression in the absence of HtrI that indicate HtrI influences proton-transfer reactions at the chromophore-attachment site essential in the transitions between the SR-I attractant signaling conformation (S₃₇₃) and the prestimulus state (SR-I₅₈₇) (37).

We thank William S. Lane for advice on preparation of protein samples for sequencing, Katherine Borkovich for her DNAsequencing protocol, and Mark Krebs, Karl Olson, Elena Spudich, and David Zacks for comments on the manuscript. This work was supported by the National Institutes of Health Grant RO1GM27750 and Office of Naval Research Grant N-00014-89-J-1629 (to J.L.S.). V.J.Y. was partially supported by the National Institutes of Health Training Grant 5T32GM07128.

- Bogomolni, R. A. & Spudich, J. L. (1982) Proc. Natl. Acad. Sci. USA 79, 1. 6250-6254.
- 2. Spudich, J. L. & Bogomolni, R. A. (1988) Annu. Rev. Biophys. Biophys. Chem. 17, 193-21
- Oesterhelt, D. & Marwan, W. (1990) Symp. Soc. Gen. Microbiol. 46, 3. 219-239.
- 4. 5.
- Tsuda, M., Nelson, B., Chang, C.-H., Govindjee, R. & Ebrey, T. G. (1985) *Biophys. J.* 47, 721–724.
 Yan, B., Takahashi, T., Johnson, R., Derguini, F., Nakanishi, K. & Spudich, J. L. (1990) *Biophys. J.* 57, 807–814.
- Yan, B., Nakanishi, K. & Spudich, J. L. (1991) Proc. Natl. Acad. Sci. 6. USA 88, 9412-9416.
- Bogomolni, R. A. & Spudich, J. L. (1987) Biophys. J. 52, 1071-1075.
- 8. Yan, B. & Spudich, J. L. (1991) Photochem. Photobiol. 54, 1023-1026.
- Spudich, J. L. & Bogomolni, R. A. (1984) Nature (London) 312, 509-513. 0
- Blanck, A., Oesterhelt, D., Ferrando, E., Schegk, E. S. & Lottspeich, F. (1989) EMBO J. 8, 3963-3971. 10.
- Spudich, E. N. & Spudich, J. L. (1982) Proc. Natl. Acad. Sci. USA 79, 11. 4308-4312.
- 12. Oesterhelt, D. & Marwan, W. (1987) J. Bacteriol. 169, 3515-3520.
- Marwan, W., Schafer, W. & Oesterhelt, D. (1990) EMBO J. 9, 355-362. 14. Schimz, A. & Hildebrand, E. (1987) Biochim. Biophys. Acta 923,
- 222-232. 15. Schimz, A., Hinsch, K.-D. & Hildebrand, E. (1989) FEBS Lett. 249,
- 59-61 16. Spudich, E. N., Hasselbacher, C. A. & Spudich, J. L. (1988) J. Bacte-
- riol. 170, 4280-4285. 17. Alam, M., Lebert, M., Oesterhelt, D. & Hazelbauer, G. L. (1989) EMBO J. 8. 631-639
- Hazelbauer, G. L. (1988) Can. J. Microbiol. 34, 466-474. 18.
- 19. Stock, J. B., Stock, A. M. & Mottonen, J. M. (1990) Nature (London) 334. 395-400
- 20. Bourret, R. B., Borkovich, K. A. & Simon, M. I. (1991) Annu. Rev. Biochem. 60, 401–441. Spudich, E. N., Takahashi, T. & Spudich, J. L. (1989) Proc. Natl. Acad.
- 21. Sci. USA 86, 7746-7750.
- 22. Sundberg, S. A., Bogomolni, R. A. & Spudich, J. L. (1985) J. Bacteriol. 164, 282-287.
- 23. Sundberg, S. A., Alam, M., Lebert, M., Spudich, J. L., Oesterhelt, D. & Hazelbauer, G. L. (1990) J. Bacteriol. 172, 2328-2335
- Alam, M. & Hazelbauer, G. L. (1991) J. Bacteriol. 173, 5837-5842. Spudich, E. N., Sundberg, S. A., Manor, D. & Spudich, J. L. (1986)
- Proteins 1, 239-246. 26. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Lab., Cold Spring Harbor, NY), 2nd Ed.
- 27. Stone, K. L., LoPresti, M. B., Crawford, J. M., DeAngelis, R. & Williams, K. R. (1989) in A Practical Guide to Protein and Peptide Purification for Microsequencing, ed. Matsudaira, P. T. (Academic, San Diego), pp. 31-47.
- Zillig, W., Palm, P., Reiter, W.-D., Gropp, F., Puhler, G. & Klenk, H.-P. 28. (1988) Eur. J. Biochem. 173, 473-482
- 29. Boyd, A., Kendall, K. & Simon, M. I. (1983) Nature (London) 301, 623-626
- 30. Rice, M. S. & Dahlquist, F. W. (1991) J. Biol. Chem. 266, 9746-9753.
- Krebs, M. P., Hauss, T., Heyn, M. P., RajBhandary, U. L. & Khorana,
 H. G. (1991) Proc. Natl. Acad. Sci. USA 88, 859-863.
 Lam, W. L. & Doolittle, W. F. (1989) Proc. Natl. Acad. Sci. USA 86, 5478-5482. 31.
- 32.
- Jones, J. G., Young, D. C. & DasSarma, S. (1991) Gene 102, 117-122. 33.
- Jihara, K. & Mukohata, Y. (1991) Arch. Biochem. Biophys. 286, 111–116.
 Lechner, J. & Sumper, M. (1987) J. Biol. Chem. 262, 9724–9729. 34.
- 35.
- von Heijne, G. (1984) EMBO J. 3, 2315-2318.
- 37. Spudich, E. N., Krebs, M. P., Khorana, H. G. & Spudich, J. L. (1992) Biophys. J. 61, 531 (abstr.).