Bioinformatics

LDH sequence

1. For creating lactate from pyruvate we needed an enzyme, called lactate-dehydrogenase (LDH). Therefore we used the database of the European Bioinformatics Institute (www.ebi.ac.uk) to find an appropriate enzyme. Since the Methylococcus capsulatus is a thermotolerant organism, we thought it would be worth searching for a thermotolerant LDH as well, so we made the search with the keywords "Idh thermotolerant".

	Services 🕺 Resea	ch 🔥 Training	 About us 			EMBL-EBI 👹 Hinxton 🔸
	BL-EBI		6			About this image: An entry from the Tara Oceans expedition in the prototype Image Data Repository, balk by scientists at the University of Datades, EMBL-EB, the University of Distribut and the University of Cambridge Learn more >
	ue Search service hel out EBI Search ≯	s you explore doze	ns of biological data resources.		Find a tool for your data analysis.	Share your scientific data with the world.
All	 Idh thermoto 	erant		Q	Find a tool >	
	Example searc	es: blast keratin bfl1				

Explore EMBL-EBI and our mission

The European Bioinformatics Institute (EMBL-EBI) shares data from life science experiments, performs basic research in computational biology and offers an extensive user training programme, supporting researchers in academia and industry. We are part of EMBL, Europe's flagship laboratory for the life sciences More about EMBL-EBI and our impact >

 Services
We provide freely available
data and bioinformatics
services to all facets of the
scientific community >

🕅 Research driven research >

👌 Training
 Research
 Training
 Industry
 ELIXIR

 We contribute to the advancement of biology
 We provide advanced bioinformatics training to scientists at all levels
 We help disseminate cutting-edge technologies to industry
 We support, as an ELIXIR node, the coordination of biological data provision through basic investigator

C Industry

& ELIXIR

edge technologies to node, the coordination of biological data provision throughout Europe >

2. Searching for the keywords "Idh thermotolerant" two results were shown. One of them was the protein sequence from the Bacillus coagulans 36D1 bacterium and the other one from the algae species Desertifilum. For using it in a bacterium we chose the bacterium enzyme.

🖀 EMBL-EBI 🔌 Services 🕺 Resea	rch 🔥 Training 🕕 About us 🔍	EMBL-EBI 🏐 Hinxton 🗸
EBI Search	Idh thermotolerant Examples: VAV_HUMAN, tyi1, Sulaten	X Q Build Query
Help & Documentation About EBI Search		Feedback
Search results for	Idh thermotolerant	
Showing 4 results out of 4 in All results		
Filter your results	Protein sequences (2 results)	
Source		
All results (4)	A0A1E5QEB6 (A0A1E5QEB6_9CYAN)	Related data v Views v
Protein sequences (2) Literature (2)	L-lactate dehydrogenase L-LDH Desertifilum sp. IPPAS B-1220 (Unreviewed)	Source: UniProtKB ID: A0A1E5QEB6_9CYAN
	G2TQM4 (G2TQM4_BACCO)	Related data - Views -
	L-lactate dehydrogenase L- LDH Bacillus coagulans 36D1 (Unreviewed)	Source: UniProtKB ID: G2TQM4_BACCO
	Literature (2 results)	
	Contributory roles of two I-lactate dehydrogenases for I-lactic acid	Related data Views
	production in thermotolerant Bacillus coagulans.	Source: MEDLINE
	Sun L, Zhang C, Lyu P, Wang Y, Wang L, Yu B	ID: 27885267

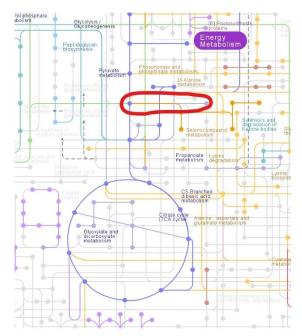
3. In the UniProt database (<u>www.uniprot.org</u>) we could find valuable information (e.g. function, names and taxonomy, structure, sequence etc.) about the chosen enzyme.

UniProt	UniProtKB +				Advanced - Q Search
BLAST Align Retrieve/ID ma	pping Peptide search			C. O	Help Conta
UniProtKB - 0	G2TQM4 (G	2TQM4_I	BACCO)		😫 Baske
Display	SBLAST = Align	Format 🔒 Add to	b basket O History	📌 Feedback 🗖 Help video	Other tutorials and vid
Entry Publications Feature viewer Feature table All N	Gene Idh Organism Bacillo		se 1 score: ●●●○○ - Protein inferred from homoloav ¹		
Function Names & Taxonomy Subcellular location	Function ⁱ Catalytic activity ⁱ				
Pathology & Biotech PTM / Processing Expression Interaction	Pathway ¹ : pyruvate This protein is involv Proteins known to be step 1. L-lactate deh This subpathway is p	involved in this subpa ydrogenase (Idh) art of the pathway pyr		s, the pathway pyruvate fermentation t	o lactate and in Fermentatio
Structure	Sites				
Family & Domains	Feature key	Position(s)	escription	Actions Graphic	al view Leng
Z Sequence	Binding site ¹	90 S	ubstrate 🖉 UniRule annotation 👻		
	Binding site ¹	122 N	AD or substrate 🗸 UniRule annotation 👻		
Similar proteins					
Similar proteins Cross-references	Binding site ¹	153 S	ubstrate 🛛 UniRule annotation 🚽		

4. For finding the gene sequence of the LDH, we visited the website of the KEGG PATHWAY Database (<u>http://www.genome.jp/kegg/pathway.html</u>), where we looked for the metabolic pathway of Bacillus coagulans.



5. In the figure we had to find the reaction of the lactate-pyruvate converting for the enzyme details.



6. Opening the enzyme's datasheet we realized that two enzymes are in connection with this reaction. We chose the L-lactate-dehydrogenase because it was smaller in size and had the usual beginning ATG bases.

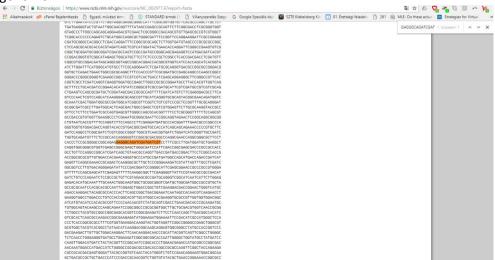
Entry	Bcoa 0653 CDS T01628	All links
Definition	(GenBank) L-lactate dehydrogenase	
ко	K00016 L-lactate dehydrogenase [EC:1.1.1.27]	Ontology (3) KEGG BRITE (3)
Organism	bag Bacillus coagulans 36D1	Pathway (8) KEGG PATHWAY (8)
Pathway	bag00010 Glycolysis / Gluconeogenesis bag00270 Cysteine and methionine metabolism bag00260 Pyruvate metabolism bag01000 Propanoate metabolism bag01100 Metabolic pathways bag01110 Biosynthesis of secondary metabolites bag01120 Microbial metabolism in diverse environments bag01130 Biosynthesis of antibiotics	Chemical substance (9) KEGG COMPOUND (° Chemical reaction (4) KEGG ENZYME (1) KEGG REACTION (3) Genome (1) KEGG GENOME (1) Gene (3) KEGG ORTHOLOGY
Brite	KEGG Orthology (KO) [BR:bag00001] Metabolism Carbohydrate metabolism 00010 Glycolysis / Gluconeogenesis Bcoa_0653 00620 Pryuvate metabolism Bcoa_0653 00640 Propanoate metabolism Bcoa_0653	NCBI-PROTEINID (1) OC (1) Protein sequence (1) UniProt (1) Protein domain (6) Plam (6) All databases (35) Download RDF
	Amino acid metabolism 00270 Cysteine and methionine metabolism Bcoa_0653 Enzymes [BR:bag01000] 1. Oxidoreductases 1.1 Acting on the CH-OH group of donors 1.1.1 With NAD+ or NADP+ as acceptor 1.1.1.27 L-lactate dehydrogenase Bcoa_0653 Exosome [BR:bag04147] Exosomal proteins Exosomal proteins	
SSDB	Bcoa_0653 Exosomal proteins of breast milk Bcoa_0653 BRITE hierarchy	
Motif	Ortholog Paralog Gene cluster GFIT Pfam: Ldh 1 N Ldh 1 C UDPG MGDP dh N ApbA 3HCDH N TrkA N	
MOUI		
Other DBs	NCBI-ProteinID: AEO99872 UniProt: GZTQM4	
Position	complement(693722694660)	
AA seq	Genome map 312 aa AA seq DB search MKKVNRIAVVGTGAVGTSYCYAMINQGVAEELVLIDINEAKAEGEAMDLNHGLPFAPTPT RVWKGDYSDCGTADLVVITAGSPOKPGETRLDLVAKNAKIFKGMIKSIMDSGFNGIFLVA SNPVDILTYVTWKESGLPKEHVIGSGTVLDSARLRNSLSAHFGIDPRNVHAAIIGEHGDT ELPVWSHTTIGYDTIESYLOKGTIDOKTLDDIFVNTRDAYHIIERKGATFYGIGMSLTR ITRAILNNENSVLTVSAFLEGQYGNSDVYIGVPAVINROGVREVVEIELNDKEQEQFSHS VKVLKETMAPVL	
NT seq	939 nt NT ang +upstream ont +downstream ont algaaaaaggtcaatogtatigcaqtggtggaacgggtgcagttggtacaagttactgc tacgccatgatatatcgagtgtggaacggggcagtggagtggagcagttggtacaagtactgc aaagcagaaggggaagccatggacctgaaccaggcctgccattgggcctacgccgacc ocgcgttggaaaggcgatattcggtatcggactggacttggtatctgtgtgcatacggcag ggtdcccgcagaaagacgtggaactaggacggacttgaggtggtattttgtgtgcaaaggaag aggcagcagtgaattaggacgaacaggcgtgacttgtgtgtcggcgaacgggaag aggcagcagtggacttggactggaccgggatttacggggattttctgtgcca agggaagcattggacgtattagtagtggaaggacgggtttaaggggattttctgtgcca agggaaccatggccgaatgtctgaccgggatttagggcgaacggg gaactccggttggggcacaagtgcttgactggattttggcacaaggagtag taccatatcattgaacggaaacaaggcgtgtaccacdgaagaggtagtctgaccg aagggaaccattgaccaaaaaacattagatattttggcacaacgggattagg taccatatcattgaacgaaaaagatggattttggtctgaccgga agggaaccattgaccaaaaaacattaggtattttggtattggacgatacag gaccaaggggtacdgaacaatgaaattaggtttgagtctcgccdtgaacgg acacaaggggaacaaggagtgtgaactggggttgaattagggtctggtttgaacgga ggccgtaaggggaacaaggagtgtgaactggggtagaacgaagaaaga	

Assembling the final construct

 It was also essential to find an appropriate promoter with high efficiency to enable the gene to work. Therefore, we looked for a promoter which can be found originally in Methylococcus capsulatus. The former iGEM team, iGEM12_juit found a nearly 1.8 kb long sequence which contained a promoter but they could not determine the exact location of the promoter.

		perience	Information	part tools			
						- F	Not Released
Part:BBa	_K730001					Regulatory	Sample Not in stock
Designed by: Antre	esh Kumar Group: iC	3EM12_juit (2012	2-09-26)				No Results
							Not Used
							Get This Part
The mxaF	gene is appro	oximately 1	.8 kb in si	ze and enco	des a 66-kDa pol	ypeptide.	
Methylotrophic ba	acteria are a diverse	group of microorg	ganisms with the	ability to utilize sin	gle-carbon (C1) substrates	more reduced than ca	arbon dioxide as their
sole source of car	rbon and energy.Met	thanotrophs poss	esses native met	thanol-inducible pr	omoters, notably promoters	which are located up	stream of genes that
					required for the synthesis o		
					tinized both biochemically a		
					r is cloned in expression ve		
					stem involves the utilization		in inducer for MxaF
		irsion of the flux t	hus leading to a	faster degradation	of methane for the cell to s	urvive.	
Sequence and Fe	aatures						
Subparts Rule	er I SS I DS	Length: 1000 b	D			View plasmid () Get part sequence.
		9					
1	100 200	300	400 500	600	700 800	900 IK	
							-
_			mx	aFpromoter			-
Assembly Compa	atibility: 10 12	24 23 25	1000				
					0.1		[ed
					Categories		
Parameters							
None							
None	arts.igem.org/cgi/pa	rtsdb/puttext.cgi	i				
None	arts.igem.org/cgi/pa • cPanel Bejelentkezés			STANDARD érmek	E [] Villanyszerelés Szego	G Google Speciális	képi 🜠 SZTE Klebelsbe
None → C ① pa Alkalmazások d	cPanel Bejelentkezés	Egyedi, művi) STANDARD érmek	E 🗋 Villanyszerelés Szege	G Google Speciális	képi 📓 SZTE Klebelsbe
None → C ① pa Alkalmazások d Ba_K730001 Part	CPanel Bejelentkezés -only sequence (1)	: 🗋 Egyedi, művi 800 bp)	észi érma 🗋 🏾 🧯			G Google Speciális	képi 👩 SZTE Klebelsbe
None → C ① pr Alkalmazások C Ba_K730001 Part aggtaccgccgttc. gccctcgaaatgcgg	 cPanel Bejelentkezés conly sequence (1) ctcttcctgaatcgcca gcgaattccggcaggta 	 Egyedi, művi 000 bp) gtgagcgttgggcca gaactccggcaggcg 	észi érme 🗋 🏽 🤅 agoggaoggtttto ggoggotgtaatca	ctcgcgggttttttcc ccgttttccaggtctt	agcgtgccgaccgtgtggttga	G Google Speciális	képi 📓 SZTE Klebelsbe
None → C ① pa Alkalmazások Ba_K730001 Part aaggtaccgccgttc: ggcctcgaaatgcgg cggatttgaacgccc	CPanel Bejelentkezés conly sequence (1) ctcttcctgaatcgcca gcgaattccggcaggta ggcccagggtggtgtgg	Egyedi, művi 000 bp) gtgagcgttgggcca gaactccggcaggc acgaccaggtacaco	észi érme 🗋 🌘 agoggoaoggtttto ggoggotgtaatoa ogtgaoggogagtg	ctcgcgggttttttcc ccgttttccaggtctt ccaccatcagcagcag	agcgtgccgaccgtgtggttga ccagccttcgaggatgatgatgc aaaccccgtgcttcgatccag	G Google Speciális	képi 📓 SZTE Klebelsbe
None Akalmazások Ba_K730001 Part aaggtaccgccgttc gccctgaaatgcgg cggatttgaacgccc tctggcgattcgatc	cPanel Bejelentkezés conly sequence (1) ctcttcctgaatcgcca gcgaattccggcaggta ggccaggtggtggcgtcaa	 Egyedi, művi Bogyadi, művi Biggagcgttgggcca ggagactccggcaggi acgaccaggtacaci cggtgatctggatca 	észi érme 🗋 🌔 agcgcacggttttce ggcggctgtaatca cgtgacggcgagtg atcgggtgccgat	ctcgcgggttttttcc ccgttttccaggtctt ccaccatcagcagcag ctggtgcagatgtttc	agcgtgccgaccgtgtggttga cccagccttcgaggatgatgcc gaaaccccgtgcttcgatccag tccgccaccagggggtccggcg	G Google Speciális	képi 👹 SZTE Klebelsbe
None → C ① pa Alkalmazások C Ba_K730001 Part aaggtaccgccgttc gccctcgaattgaacgccc ctcggcattgaatgcgaa ctcaggtagcgaactcgtc acggcccaggcgaatt	CPanel Bejelentkezés conly sequence (1) ctcttcctgaatcgcca gcgaattccggcaggta ggcccagggtggtggtgg ggccgggttggcgtcaa caggcggggcgtcatcaa caggcggggcgg	Egyedi, műv 800 bp) gtgagcgttgggcca gaactccggcaggcg acgatcaggtacac cggtgatctggatca accetcgcggggc tgggcgatcattcg	észi érme [] () agcgcacggttttce ggcggctgtaatca- cgtgacggcgagtg- atcggstgccgagtg- cggcagagtgcgagtg- gaccggcgagcgagcgac	ctcgcgggttttttcc ccgttttccaggtctt ccaccatcagcagcag ctggtgcagatgttc agatcgatgatcgtcg ccgccgccaccgcctg	agcgtgccgaccgtgtggttga ccagccttcgaggatgatgc aaaccccgtgcttcgatccag tccgccaccagggggtccggcg tttcgccttgatggatgctgg tttccagccggcatcgatcagc	G Google Speciális	kép: 📓 SZTE Klebelsbe
None Alkalmazások Ba_K730001 Part aaggtaccgcgttc ggatttgaacgccg ccccggaattgaagccg cccggattgacggac tcaggtsgcsggtgac	CPanel Bejelentkezés -only sequence (1) cctcttcctgcaatggca gggaattcggcaggta ggcccaggtggtggtgg ggccgggtggtgggta caggcgggcgcttcctc ggtcgagccgggcgcage cgatgaccgggtcgattct	Byedi, műv 000 bp) gtgagcgttgggcci gaactccggcaggci cggtgatctggatca accctccgcggggcc tgggcaccatcg	észi érme 🗋 🤅 agcgcacggttttc ggcggctgtaatca cgggaggggag	ctcgcgggttttttcc ccgttttccaggtctt ccaccatcagcagcag ctggtgcagatgtttc agatcggtgcagatgtttc ccgccgccaccgcctg ccgcagcaccggtgccc	agcgtgccgaccgtgtggttga ccagccttcgaggatgatgc aaaccccgtgcttcgatcag ttcgccaccagggggtcggc tttcgcctgaggggtgctgag tttcaggcggatggtgatggc ttcaggcaggatggcaga	G Google Speciális	képi 🐻 SZTE Klebelsbe
None Alkalmazások Sa_K730001 Part asgstaccgccgttc ccctgagaatcgag ggattgaacgccc tcggcgattgacgcc tcggcgattgac tcggcgattgac asgccaggtgac asgccaggtgac asgtcatccgacgg	• cPanel Bejelentkezés -only sequence (1) cttttctgaatcgca ggccagttcggcagtaggcgaggcg	Byedi, művi Bygacgttgggcci gagactcaggtaggci acgaccaggtacaci ccggtgatctggatc accetccgcggggci ccggcaccagscgg acceacgacgacg acceacgacgacg acceacgacgacg acceacgacgacg acceacgacgacg acceacgacgacg acceacgacgacg acceacgacgacg acceacgacgacg acceacgacgacgacgacg acceacgacgacgacg acceacgacgacgacg acceacgacgacgacgacg acceacgacgacgacgacg acceacgacgacgacgacgacg acceacgacgacgacgacgacgacgacgacgacgacgacgacga	észi érme 🗋 🤅 agegecaeggetttee ggeggetgtaatea egtgaeggegggggg ategggtgeggg gaeeggegggeggae ggeeggegggggae geteceggaageggae geteceggaagegae	ctcgrgggttttttcc ccgttttccaggtctt ccaccatcagcagcag ttggtgcagatgttc ccgrcgccaccgtcg ccgrcgccaccgctg cacgaacagggtgccc tcgtattagtttgcct gcagacattcgaagt	agcgtgccgaccgtgtggttga ccagccttcgaggatgatgc aaaccccgtgcttcgatccag tccgccaccagggggtccggcg tttcgccttgatggatgctgg tttccagccggcatcgatcagc	G Google Speciális	kėpi 🔞 SZTE Klebelsbe

2. We did a search to find the 1.8 kb long sequence in Methylococcus capsulatus' complete genome.



C Biztonságos https://www.ncbi.nlm.nih.gov/nuccore/NC		_			@ ☆
III Alkalmazások 🕐 cPanel Bejelentkezés 🗅 Egyedi, művészi érme 🗅 🏐 S NCBI Resources 🕑 How To 😒	STANDARD érmek E 📋 Villanyszerelés Szege G	Google Speciális képi 📓	SZTE Klebelsberg Kör	01. Érettségi felac	iats: 🗋 261 🎯 MLE-Dothes
Nucleotide Nucleotide Advanced					Search
Graphics - Methylococcus capsulatus str. Bath, compl NCBI Reference Sequence: NC_002977.6	ete genome			Send to: -	
GenBank FASTA	Search Results			A X dback	*
<u>и 1500 к. 1400 к. 1600 к. 1600 к. 11 М. 17500 к.</u>	Features Components Sequence Th Label ^ GAGGGCAGATCGATGATCGT	Prom 822734		Strand	Ger
NC_002977.6:1.7HL.TH (49bp) Find: ✓ ✓ ↓ K52.210 ↓ K52.210 <t< td=""><td>GRUGGLIGAT LGT GAT LGT</td><td>822/34</td><td>822733</td><td>228 6 6 7 7 7 6 7 7 7 7 7 7 7 7 7 7 7 7 7</td><td>2017 Ameri Fi</td></t<>	GRUGGLIGAT LGT GAT LGT	822/34	822733	228 6 6 7 7 7 6 7 7 7 7 7 7 7 7 7 7 7 7 7	2017 Ameri Fi
ST3 Marke [®] [©] [©] [©] ^K ^T Repest region L652200 1.652290			Displaying Search Re	328	Analyze this sequence Run BLAST Pick Primers Highlight Sequence Features
					Related information Assembly
					BioProject
					BioSample
					Components (Core)

3. This sequence contained not only the nucleotide sequence between two genes (the moxY and the mxaF) but did contain a partial part of each of the two genes.

200 K		1,200 K	1,400 K 1,600 K 1,800 K	2 M	2,200		2,600 K	2,800 K	3 M
NC_002977.6: 823K	823K (47bp) - Find:	~ (\$\$) Q -	// Q, 🚮 🗮				>	🕈 Tools 🔹 🚠	🏟 Tracks <i>@</i>
823,020	823,838		823,848	8	323,850			823,868	
іс в с т т	т в с т с с с в в в	AAAGA	TCGTATT	AGTT	ΤG	сстс	GAT	CGG	CGG
GCGA) nes	ACGAGGCCC	тттст	A G C A T A A	ТСАА	A C	GGAG	CTA	всс	всс
75 Markers peat region 223,028	Control C	ensor 0)	LL 2017 12.000 August 204023047 21.2000 Change 21.2 Modify Range 22.3 Modify Range 22.4 Modify Range 22.4 Modify Range 23.4 Modify Range 24.4 Modify Range 25.4 Modify Celection 20.2 Modify Range 25.4 Modify Celection 20.2 Modify Range 25.4 Modify Celection 26.4 Modify Celection 27.4 Modify Range 27.4 Modify Celection 27.4		323,850	methan Location: 82,27 [Length] Span: 1,806 Product: 601 [Qualifiers] inference: COORC sequer Download: <u>WP 03</u> Links & Tools BLAST Genomic: BLAST Genomic:	0960106.1 ependent dehyd nol(ethanol fami 8825,083 DINATES: simila cce:RefSeq:WP_ 10960106.1 WP_010960106 WP_010960106 WP_010960106 WP_010960106	r to AA 010960106.1 823,278825,0 1.1 1.1 823,278825,0 823,278825,0	83), WP 010960

4. The putative moxY or mxaF used in this study did only contain the nucleotide sequence between the two genes, which must contain the promoter of either the moxY or the mxaF gene. Unfortunately, the orientation of the promoter is not known because the moxY and the mxaF genes are in different directions, therefore we could not determine neither the exact orientation of the promoter nor the exact sequence but we managed to approach the exact sequence of the promoter and apply it in such a way that the orientation was not needed to know. 5. Having the sequence of the LDH gene and an appropriate promoter, only a usable vector was needed to transfer the gene and the promoter into Methylococcus capsulatus. An article wrote about the vectors pMHE2, pMHE3, pMHE5, pMHE6 and pMHE7 (figure). We chose the pMHE5 and pMHE7 vectors because they were available for us in short time.

	T7 promoter prim	er					stream	127-15 BEECH & R	c 1900,000,000
Bglll	T7 promoter	lac operator	<u>Xbal</u>	RBS	Ncol		<u>Stul Hin</u> ctor-Xa Sal	I Clai	EcoRV Pa
Xmal Smal BamHI	SCTANCAASCCCGAAASGAA			T7 terminator	MetGly#	lisHisHisHisHisHisIl			
InProSlySlySerSlyC			T7 terminator prim		TCTTGRUGGG	TTTTTGC			pMHE2
	T7 promoter prim	ier							
<u>Bg/II</u>	T7 promoter	lac operator	Xbal	RBS	<u>Ndel</u>	Т7-tag	BamHI		HindIII
1202 S2.12000					MetAla	SerMetThrSlySlySlmG	lnNetGlyArgAspPro	AanderderderValAap	LysLeuAlaAlaA
Khol 6His-tag	NCCNCTGNGATCCSSCTSCTA	ACAMAGCCOGAMAGGAAGCTGA	GTTGGCTGCTGCCACCGCTGA	CAATAACTAGCATAACCOC	T7 term	inator www.seetcttelesecttt	PPPoc	-	
/ev31uRisHisHisHis	lisHis		T7 t	erminator primer				pMHE3, p	MHE3Te
	T7 promoter prim	her		78: 				98.00 (199.00) 198.00 (199.00)	a
<u>Bg/II</u>	T7 promoter	lac operator	Xbal	RBS	Ncol		<u>Stul Hir</u> ctor-Xa Sa	I* Clal	EcoRV* I
Marterearreeoodaa			TELECTETAGAAATAATTITIS	111/00/1112/00/00/00/00/00/00/00/00/00/00/00/00/00		ATCATCATCATCATCATAT IIshishishishishisii			
Xhol BamHI*	FLAG-tag	Strep-tag II	GAAGTGAAGATCOGGCTGCTA	ACAAAGCCCGAAAGGAAGCT	MOTTOOCTOC	NOCACOSCI SASCAATAAC		7 terminator	AGGGGTTTTTTGC
lnLeuGluGlySerAap1	enterokinase	an7rpSerHisProGlnPheGl	ulys			T7 termina	tor primer	pMHE5, p	MHE5Te
	T7 promoter prim	ier						1999 - 1999 -	
Bg/III	T7 promoter	lac operator	Xbal	RBS	Ndel	T7-tag	ANTOOTO GOOMTOO	Sacl Sall*	HindIII
Pstl Xhol						SerMetThrGlyGlyGlnG		AanSerSerSerValAsp	LysLeuAspIleG
TOCTOCASCTOGASSGAT	FLAG-tag	Strep-tag	STTOSAGAASTGAAGATCOSS	TOTANCANASCCCGANAS	AASCTGASTTC	остостосслоодствлас	WINKTAGONIAACCO	T7 terminato	
heLeuGlnLeuGluGly5 7	enterokina		nPheGluLys			T7 to	erminator prime	pMHE6, p	MHE6Tc
	T7 promoter prim	ner							
<u>Bg/II</u>	T7 promoter	lac operator	Xbal	RBS	Ncol	FLAG-tag	Xhol	Strep-tag II	factor-Xa
Stul <u>Hinc</u> ll* <u>Sall*</u> Cle	<u>HindI</u> II <u>EcoR</u> I <u>1 EcoR</u> V* <u>I</u>	<u>Xmal</u> Pstl Smal BamHI*			MetAspA	enterokina	se T7 termin	ator	iulysIleGluGi
2000TGTOGAOSGTATOS	ATAAGCTTGATATOGAATTOC	NGCAGOCOGGGGGGATCOGGCTG suGlnProGlyGlySerGlyCy		TGASTTOSCTOCTOCACCO	CTONSCANTA	CTASCATAACCCCTTGGGG	CTCTAMOSSSTCTTG	GOOGTTTTTTTGC	MHE7Te

Figure

6. The LDH gene was intended to ligate between the Ncol and EcoRV restriction sites, therefore, it was supplied with an Ncol restriction site and an addition AGTCAGTC nucleotide sequence before and after the Ncol restriction site, in order to minimalize the possible damage made by the restriction enzymes (figure).

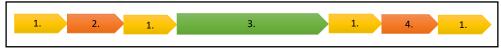


Figure: The basic structure of the synthesized gene

- 1. Addition AGTCAGTC bases
- Ncol restriction site
 LDH gene
- 4. EcoRV restriction site

7.

The LDH gene contained an Ncol restriction site in itself, therefore, we replaced the bases so that the coded amino-acid did not changed.

1 10 Ncol (20 23	30	40	50	60
		r g a a a a a g g t		GCAGTGGTTG	GAACGGGT
AGTCAGTC Ncol restri	AGTCAGTC		LDH gene		
70	80 80	90	100	110	120
GCAGTTGGTAC	ААGTTАСТGСТАС		аатсабббтб	TTGCAGAAGA	осттоттт
2		LDH gene			
130	140	150	160	170	180
<u>Clal (127)</u> TAATČG _A ATATT	A A C G A A G C A A A A (5 C A G A A G G G G	Ncol (160)	CCTGAACCAC	бесстесс
>		LDH gene			
190	200	210	220	230	240
ATTTGCGCCTA	CGCCGACCCGCG	TTTGGAAAGG	GCGATTATTCC	GATTGCGGCA	CTGCCGAT
>		LDH gene			
1 10	20	30	40	50	60
Ncol					
			CAATCGTATT		
A G T C A G T C C C A AGTCAGTC Ncol restr	(10) T G_G A G T C A G T C A AGTCAGTC	т <u>а а а а а а а а </u> а а т ⁹⁹ с а с с а т а а т т	CAATCGTATT LDH gene	G C A G T G G T T G G	5 A A C G G G T
A G T C A G T C C C A AGTCAGTC Ncol restr	(10) T G G A G T C A G T C A AGTCAGTC iction site 80	90	CAATCGTATT LDH gene	G C A G T G G T T G G	5 A A C G G G T
A G T C A G T C C C C A AGTCAGTC Ncol restr 70 G C A G T T G G T A C	(10) T G G A G T C A G T C A AGTCAGTC iction site 80	т <u>а а а а а а а а </u> а а т ⁹⁹ с а с с а т <u>а а</u> т т	CAATCGTATT LDHgene	G C A G T G G T T G G	5 A A C G G G T
A G T C A G T C C [*] C A AGTCAGTC Ncol restr 76 G C A G T T G G T A C	(10) T G_G A G T C A G T C A AGTCAGTC iction site A A G T T A C T G C T A G	F G A A A A A A G G T 90 E G C C A T G A T T LDH gene 150 5 C A G A A G G G G	CAATCGTATT LDHgene	G C A G T G G T T G C 110 T T G C A G A A G A C	120 5 C T T G T T T 180
A G T C A G T C C [*] C A AGTCAGTC Ncol restr 76 G C A G T T G G T A C	(10) T G G A G T C A G T C A T A AGTCAGTC 50 A A G T T A C T G C T A C 140	50 50 50 50 50 50 50 50 50 50 50 50	CAATCGTATT LDHgene	G C A G T G G T T G C 110 T T G C A G A A G A C	120 5 C T T G T T T 180
A G T C A G T C C [*] C A AGTCAGTC Ncol restr 76 G C A G T T G G T A C	(10) T G G A G T C A G T C A T A AGTCAGTC 50 A A G T T A C T G C T A C 140	F G A A A A A A G G T 90 E G C C A T G A T T LDH gene 150 5 C A G A A G G G G	CAATCGTATT LDHgene	G C A G T G G T T G C 110 T T G C A G A A G A C	120 5 C T T G T T T 180
A G T C A G T C C C A AGTCAGTC Ncol restr 76 G C A G T T G G T A C 130 T A A T C G A T A T T 180	(10) T G G A G T C A G T C A G AGTCAGTC iction site 50 A A G T T A C T G C T A C 140 140 A A C G A A G C A A A A C	F G A A A A A A G G T Si C G C C A T G A T T LDH gene 150 5 C A G A A G G G G LDH gene 210 T T T G G A A A G G		G C A G T G G T T G C 110 T T G C A G A A G A G 170 C C T G A A C C A C C 230	120 G C T T G T T T 180 G C C T G C C 240
A G T C A G T C C C A AGTCAGTC Ncol restr 76 G C A G T T G G T A C 130 T A A T C G A T A T T 180	(10) T G G A G T C A G T C A G AGTCAGTC iction site A A G T T A C T G C T A G 140 A A C G A A G C A A A A A G 200	S S S S S C C C C C A C A A A A A A A G G C C A T A A A A A G G T A A A A A G T A A A A G T A C T A A T T LDH gene S C C A A A A A C A T T LDH gene S C A A A A A A A A A A A A A		G C A G T G G T T G C 110 T T G C A G A A G A G 170 C C T G A A C C A C C 230	120 G C T T G T T T 180 G C C T G C C 240

8. The promoter was intended to ligate between the BgIII restriction sites. This could enable the promoter to ligate in both orientations, randomly. The promoter was supplied with BgIII restriction sites at both ends and addition AGTCAGTC nucleotides before and after each added AGTCAGTC nucleotides (figure).

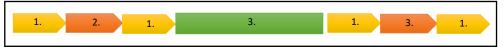


Figure: The basic structure of the synthesized promoter

- 1. Addition AGTCAGTC bases
- 2. Bglll restriction site
- 3. Promoter