

SEQUENCE LISTING

<110> Cornell University
Adverum Biotechnologies

<120> GENE THERAPY FOR THE TREATMENT OF ALDEHYDE DEHYDROGENASE
DEFICIENCY

<130> 725826

<160> 5

<170> PatentIn version 3.5

<210> 1
<211> 500
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic mature 500 aa protein

<400> 1

Ser Ala Ala Ala Thr Gln Ala Val Pro Ala Pro Asn Gln Gln Pro Glu
1 5 10 15

Val Phe Cys Asn Gln Ile Phe Ile Asn Asn Glu Trp His Asp Ala Val
20 25 30

Ser Arg Lys Thr Phe Pro Thr Val Asn Pro Ser Thr Gly Glu Val Ile
35 40 45

Cys Gln Val Ala Glu Gly Asp Lys Glu Asp Val Asp Lys Ala Val Lys
50 55 60

Ala Ala Arg Ala Ala Phe Gln Leu Gly Ser Pro Trp Arg Arg Met Asp
65 70 75 80

Ala Ser His Arg Gly Arg Leu Leu Asn Arg Leu Ala Asp Leu Ile Glu
85 90 95

Arg Asp Arg Thr Tyr Leu Ala Ala Leu Glu Thr Leu Asp Asn Gly Lys
100 105 110

Pro Tyr Val Ile Ser Tyr Leu Val Asp Leu Asp Met Val Leu Lys Cys

115	120	125													
Leu	Arg	Tyr	Tyr	Ala	Gly	Trp	Ala	Asp	Lys	Tyr	His	Gly	Lys	Thr	Ile
130						135					140				
Pro	Ile	Asp	Gly	Asp	Phe	Phe	Ser	Tyr	Thr	Arg	His	Glu	Pro	Val	Gly
145					150					155					160
Val	Cys	Gly	Gln	Ile	Ile	Pro	Trp	Asn	Phe	Pro	Leu	Leu	Met	Gln	Ala
				165					170					175	
Trp	Lys	Leu	Gly	Pro	Ala	Leu	Ala	Thr	Gly	Asn	Val	Val	Val	Met	Lys
			180					185					190		
Val	Ala	Glu	Gln	Thr	Pro	Leu	Thr	Ala	Leu	Tyr	Val	Ala	Asn	Leu	Ile
		195					200					205			
Lys	Glu	Ala	Gly	Phe	Pro	Pro	Gly	Val	Val	Asn	Ile	Val	Pro	Gly	Phe
	210					215					220				
Gly	Pro	Thr	Ala	Gly	Ala	Ala	Ile	Ala	Ser	His	Glu	Asp	Val	Asp	Lys
225					230					235					240
Val	Ala	Phe	Thr	Gly	Ser	Thr	Glu	Ile	Gly	Arg	Val	Ile	Gln	Val	Ala
				245					250					255	
Ala	Gly	Ser	Ser	Asn	Leu	Lys	Arg	Val	Thr	Leu	Glu	Leu	Gly	Gly	Lys
			260					265					270		
Ser	Pro	Asn	Ile	Ile	Met	Ser	Asp	Ala	Asp	Met	Asp	Trp	Ala	Val	Glu
		275					280					285			
Gln	Ala	His	Phe	Ala	Leu	Phe	Phe	Asn	Gln	Gly	Gln	Cys	Cys	Cys	Ala
	290					295					300				
Gly	Ser	Arg	Thr	Phe	Val	Gln	Glu	Asp	Ile	Tyr	Asp	Glu	Phe	Val	Glu
305					310					315					320
Arg	Ser	Val	Ala	Arg	Ala	Lys	Ser	Arg	Val	Val	Gly	Asn	Pro	Phe	Asp
				325					330					335	

Ser Lys Thr Glu Gln Gly Pro Gln Val Asp Glu Thr Gln Phe Lys Lys
340 345 350

Ile Leu Gly Tyr Ile Asn Thr Gly Lys Gln Glu Gly Ala Lys Leu Leu
355 360 365

Cys Gly Gly Gly Ile Ala Ala Asp Arg Gly Tyr Phe Ile Gln Pro Thr
370 375 380

Val Phe Gly Asp Val Gln Asp Gly Met Thr Ile Ala Lys Glu Glu Ile
385 390 395 400

Phe Gly Pro Val Met Gln Ile Leu Lys Phe Lys Thr Ile Glu Glu Val
405 410 415

Val Gly Arg Ala Asn Asn Ser Thr Tyr Gly Leu Ala Ala Ala Val Phe
420 425 430

Thr Lys Asp Leu Asp Lys Ala Asn Tyr Leu Ser Gln Ala Leu Gln Ala
435 440 445

Gly Thr Val Trp Val Asn Cys Tyr Asp Val Phe Gly Ala Gln Ser Pro
450 455 460

Phe Gly Gly Tyr Lys Met Ser Gly Ser Gly Arg Glu Leu Gly Glu Tyr
465 470 475 480

Gly Leu Gln Ala Tyr Thr Glu Val Lys Thr Val Thr Val Lys Val Pro
485 490 495

Gln Lys Asn Ser
500

<210> 2

<211> 1503

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic nucleotide sequence encoding mature 500 aa

protein

<400> 2

tcagccgccg ccacccaggc cgtgcctgcc cccaaccagc agcccgaggt cttctgcaac	60
cagattttca taaacaatga atggcacgat gccgtcagca ggaaaacatt cccaccgctc	120
aatccgtcca ctggagaggt catctgtcag gtagctgaag gggacaagga agatgtggac	180
aaggcagtga aggccgcccc ggccgccttc cagctgggct caccttggcg ccgcatggac	240
gcatcacaca ggggccggct gctgaaccgc ctggccgata tgatcgagcg ggaccggacc	300
tacctggcgg ccttggagac cctggacaat ggcaagccct atgtcatctc ctacctggtg	360
gatttggaca tggtcctcaa atgtctccgg tattatgccg gctgggctga taagtaccac	420
gggaaaacca tccccattga cggagacttc ttcagctaca cacgccatga acctgtgggg	480
gtgtgcgggc agatcattcc gtggaatttc ccgctcctga tgcaagcatg gaagctgggc	540
ccagccttgg caactggaaa cgtgggttgtg atgaaggtag ctgagcagac acccctcacc	600
gccctctatg tggccaacct gatcaaggag gctggctttc cccctggtgt ggtcaacatt	660
gtgcctggat ttggccccac ggctggggcc gccattgcct cccatgagga tgtggacaaa	720
gtggcattca caggctccac tgagattggc cgcgtaatcc aggttgctgc tgggagcagc	780
aacctcaaga gagtacctt ggagctgggg gggaagagcc ccaacatcat catgtcagat	840
gccgatatgg attgggccgt ggaacaggcc cacttcgccc tgttcttcaa ccagggccag	900
tgctgctgtg ccggctcccc gaccttcgtg caggaggaca tctatgatga gtttgtggag	960
cggagcgttg cccgggcca gtctcgggtg gtcgggaacc cctttgatag caagaccgag	1020
caggggccgc aggtggatga aactcagttt aagaagatcc tcggctacat caacacgggg	1080
aagcaagagg gggcgaagct gctgtgtggt gggggcattg ctgctgaccg tggttacttc	1140
atccagccca ctgtgttttg agatgtgcag gatggcatga ccatcgccaa ggaggagatc	1200
ttcgggccag tgatgcagat cctgaagttc aagaccatag aggaggttgt tgggagagcc	1260
aacaattcca cgtacgggct ggccgcagct gtcttcacaa aggatttgga caaggccaat	1320
tacctgtccc aggccctcca ggcgggcact gtgtgggtca actgctatga tgtgttttga	1380
gccagtcac cctttggtgg ctacaagatg tcggggagtg gccgggagtt gggcgagtac	1440
gggctgcagg catacactga agtgaaaact gtcacagtca aagtcctca gaagaactca	1500

taa

1503

<210> 3

<211> 517

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic immature 517 aa protein including signal
sequence

<400> 3

Met Leu Arg Ala Ala Ala Arg Phe Gly Pro Arg Leu Gly Arg Arg Leu
1 5 10 15

Leu Ser Ala Ala Ala Thr Gln Ala Val Pro Ala Pro Asn Gln Gln Pro
20 25 30

Glu Val Phe Cys Asn Gln Ile Phe Ile Asn Asn Glu Trp His Asp Ala
35 40 45

Val Ser Arg Lys Thr Phe Pro Thr Val Asn Pro Ser Thr Gly Glu Val
50 55 60

Ile Cys Gln Val Ala Glu Gly Asp Lys Glu Asp Val Asp Lys Ala Val
65 70 75 80

Lys Ala Ala Arg Ala Ala Phe Gln Leu Gly Ser Pro Trp Arg Arg Met
85 90 95

Asp Ala Ser His Arg Gly Arg Leu Leu Asn Arg Leu Ala Asp Leu Ile
100 105 110

Glu Arg Asp Arg Thr Tyr Leu Ala Ala Leu Glu Thr Leu Asp Asn Gly
115 120 125

Lys Pro Tyr Val Ile Ser Tyr Leu Val Asp Leu Asp Met Val Leu Lys
130 135 140

Cys Leu Arg Tyr Tyr Ala Gly Trp Ala Asp Lys Tyr His Gly Lys Thr
145 150 155 160

Ile Pro Ile Asp Gly Asp Phe Phe Ser Tyr Thr Arg His Glu Pro Val
165 170 175

Gly Val Cys Gly Gln Ile Ile Pro Trp Asn Phe Pro Leu Leu Met Gln
180 185 190

Ala Trp Lys Leu Gly Pro Ala Leu Ala Thr Gly Asn Val Val Val Met
195 200 205

Lys Val Ala Glu Gln Thr Pro Leu Thr Ala Leu Tyr Val Ala Asn Leu
210 215 220

Ile Lys Glu Ala Gly Phe Pro Pro Gly Val Val Asn Ile Val Pro Gly
225 230 235 240

Phe Gly Pro Thr Ala Gly Ala Ala Ile Ala Ser His Glu Asp Val Asp
245 250 255

Lys Val Ala Phe Thr Gly Ser Thr Glu Ile Gly Arg Val Ile Gln Val
260 265 270

Ala Ala Gly Ser Ser Asn Leu Lys Arg Val Thr Leu Glu Leu Gly Gly
275 280 285

Lys Ser Pro Asn Ile Ile Met Ser Asp Ala Asp Met Asp Trp Ala Val
290 295 300

Glu Gln Ala His Phe Ala Leu Phe Phe Asn Gln Gly Gln Cys Cys Cys
305 310 315 320

Ala Gly Ser Arg Thr Phe Val Gln Glu Asp Ile Tyr Asp Glu Phe Val
325 330 335

Glu Arg Ser Val Ala Arg Ala Lys Ser Arg Val Val Gly Asn Pro Phe
340 345 350

Asp Ser Lys Thr Glu Gln Gly Pro Gln Val Asp Glu Thr Gln Phe Lys
355 360 365

Lys Ile Leu Gly Tyr Ile Asn Thr Gly Lys Gln Glu Gly Ala Lys Leu
 370 375 380

Leu Cys Gly Gly Gly Ile Ala Ala Asp Arg Gly Tyr Phe Ile Gln Pro
 385 390 395 400

Thr Val Phe Gly Asp Val Gln Asp Gly Met Thr Ile Ala Lys Glu Glu
 405 410 415

Ile Phe Gly Pro Val Met Gln Ile Leu Lys Phe Lys Thr Ile Glu Glu
 420 425 430

Val Val Gly Arg Ala Asn Asn Ser Thr Tyr Gly Leu Ala Ala Ala Val
 435 440 445

Phe Thr Lys Asp Leu Asp Lys Ala Asn Tyr Leu Ser Gln Ala Leu Gln
 450 455 460

Ala Gly Thr Val Trp Val Asn Cys Tyr Asp Val Phe Gly Ala Gln Ser
 465 470 475 480

Pro Phe Gly Gly Tyr Lys Met Ser Gly Ser Gly Arg Glu Leu Gly Glu
 485 490 495

Tyr Gly Leu Gln Ala Tyr Thr Glu Val Lys Thr Val Thr Val Lys Val
 500 505 510

Pro Gln Lys Asn Ser
 515

<210> 4
 <211> 1554
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic nucleic acid encoding immature 517 aa protein
 including signal sequence

<400> 4
 atgttgcgctgctgccccgcttcggggcccgcctggggccgccgcctcttgtcagccgcc 60
 gccacccaggccgtgcctgc cccaaccagcagcccagggtcttctgcaa ccagattttc 120

ataaacaatg aatggcacga tgccgtcagc aggaaaacat tccccaccgt caatccgtcc	180
actggagagg tcatctgtca ggtagctgaa ggggacaagg aagatgtgga caaggcagtg	240
aaggccgccc gggccgcctt ccagctgggc tcaccttggc gccgcatgga cgcatcacac	300
aggggccggc tgctgaaccg cctggccgat ctgacgagc gggaccggac ctacctggcg	360
gccttggaga ccctggacaa tggcaagccc tatgtcatct cctacctggt ggatttggac	420
atggtcctca aatgtctccg gtattatgcc ggctgggctg ataagtacca cgggaaaacc	480
atccccattg acggagactt cttcagctac acacgccatg aacctgtggg ggtgtgcggg	540
cagatcattc cgtggaattt cccgctcctg atgcaagcat ggaagctggg cccagccttg	600
gcaactggaa acgtggttgt gatgaaggta gctgagcaga caccctcac cgccctctat	660
gtggccaacc tgatcaagga ggctggcttt ccccctggtg tggtaacat tgtgcctgga	720
tttggcccca cggctggggc cgccattgcc tcccatgagg atgtggacaa agtggcattc	780
acaggctcca ctgagattgg ccgcgtaatc caggttgctg ctgggagcag caacctcaag	840
agagtgacct tggagctggg ggggaagagc cccaacatca tcatgtcaga tgccgatatg	900
gattgggccc tggaacaggc ccacttcgcc ctgttcttca accagggccca gtgctgctgt	960
gccggctccc ggaccttcgt gcaggaggac atctatgatg agtttgtgga gcggagcgtt	1020
gccccgggcca agtctcgggt ggtcgggaac ccctttgata gcaagaccga gcaggggccc	1080
cagggtggatg aaactcagtt taagaagatc ctcggctaca tcaacacggg gaagcaagag	1140
ggggcgaagc tgctgtgtgg tgggggcatt gctgctgacc gtggttactt catccagccc	1200
actgtgtttg gagatgtgca ggatggcatg accatcgcca aggaggagat cttcgggcca	1260
gtgatgcaga tcctgaagtt caagaccata gaggaggttg ttgggagagc caacaattcc	1320
acgtacgggc tggccgcagc tgtcttcaca aaggatttgg acaaggccaa ttacctgtcc	1380
caggccctcc aggcgggcac tgtgtgggtc aactgctatg atgtgtttgg agcccagtca	1440
ccctttggtg gctacaagat gtcggggagt ggccgggagt tgggcgagta cgggctgcag	1500
gcatacactg aagtgaaaac tgtcacagtc aaagtgcttc agaagaactc ataa	1554

<210> 5
 <211> 17
 <212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Signal peptide by itself

<400> 5

Met	Leu	Arg	Ala	Ala	Ala	Arg	Phe	Gly	Pro	Arg	Leu	Gly	Arg	Arg	Leu
1				5					10					15	

Leu