ORIGINAL PAPER

# Structure of Patt1 human proapoptotic histone acetyltransferase

Roch Paweł Jędrzejewski · Rajmund Kaźmierkiewicz

Received: 27 July 2013 / Accepted: 21 October 2013 / Published online: 19 November 2013 © The Author(s) 2013. This article is published with open access at Springerlink.com

Abstract The results of modeling of a novel human histone acetyltransferase Patt1 are presented here. This protein belongs to the GNAT GCN5 family and shows proapoptotic activity in human hepatocellular carcinoma cells. Patt1 is an attractive therapeutic target. The sequence analysis, fold recognition predictions and homology modeling of Patt1 protein structure were performed. N- and C- termini of Patt1 were unstructured. Central part revealed classical GNAT fold-central 7-stranded beta sheet core surrounded by intervening 4 alpha helices. The model was assessed with the methods for protein structure validation PROQ and MetaMQAPII. The allatom 12 ns molecular dynamics simulation of Patt1 model with TIP3P water model and counterions was conducted. All assessment methods implemented resulted in conviction that the model was of quality that could provide confident structural information to infer sequence-structure-function relationships of Patt1. Phe186 and Cys137 were identified as residues engaged in acetyltransfer reaction and the clues for the identification of reaction mechanism were proposed. The knowledge of detailed molecular architecture of Patt1 is not only the key to understanding its mechanistic functional properties but it also opens the possibility of rational drug and protein design experiments, leading to development of effective therapeutic methods.

**Keywords** GNAT family proteins · HAT histone acetyltransferases · Homology modeling · Molecular dynamics simulation · Protein structure prediction

R. P. Jędrzejewski (🖂) · R. Kaźmierkiewicz

# Introduction

In eukaryotic cells the acetylation of histone N-terminal tails is the key regulatory mechanism in histone code and the execution of epigenetic information [1-3]. The information encoded in modification patterns of N-terminal histone tails is crucial for chromatin remodeling. The modification of tails is then read out by specific proteins by means of molecular recognition. Chromatin remodeling machines (CRM) complexes take the input information encoded in the form of the modifications order. This drives the chromatin remodeling processes and opens access to chromatin. These events transfer chromatin to transcriptionally active state in which DNA template is accessible for transcription factors and other regulatory proteins that bind DNA. This results in expression of specific genes. Switching chromatin to an open state is induced by acetylation [1–3]. Acetylation is performed by vast array of enzymes from various families [4]. Many different families of proteins are engaged in histone acetylation [5, 6]. Histone acetyltransferases (HAT) perform acetylation of N-terminal histone tails at specific lysines. Dysfunction of histone acetyltransferases leads to carcinogenic processes [7, 8]. This aspect of acetyltransferases has gained growing attention in recent years. The GNAT family of acetyltransferases [9-11] contains the first discovered proteins that possess histone acetylation activity that was connected with gene activation [12]. A recently discovered member of this group Patt1 (protein acetyltransferase-1) has been shown to acetylate histone H4 in vitro and in vivo [13]. Activity of this protein was linked to promotion of apoptosis in human hepatocellular carcinoma cells. There is no information available about the structure of this new human histone acetyltransferase. Patt1 protein allatom tertiary structure was modeled with the use of theoretical methods. All structure assessment methods and the results of molecular dynamics simulation indicated that the model is correct. The model exhibits all features of GNAT canonical

Laboratory of Biomolecular Systems Simulations, Intercollegiate Faculty of Biotechnology, University of Gdańsk and Medical University of Gdańsk, Kładki 24, 80-822 Gdańsk, Poland e-mail: rochjedrzejewski@gmail.com

topology: the core of protein is formed by central beta sheet composed of seven strands. Alpha helices are mixed with beta sheets and surround the structure on both sides. The enzyme active site is located at the edge of beta sheet, between  $\beta 5$  and  $\beta 6$  structural elements. Amino acid residues engaged in acetyltransfer reaction were identified: the aromatic rings of Phe185, Phe186, and Phe192 form a stacking system. Their role is alternative to general acid catalysis. The spatial localization of Cys137 is similar to analogous cysteine residue in Esa1 histone acetyltransferase where acetyltransfer reaction proceeds through an acetyl-cysteine enzyme intermediate. The model represents the first structural data regarding Patt1 protein.

## Materials and methods

### Patt1 sequence analysis

The searches of an nr National Center for Biotechnology Information (NCBI) sequence database [14] were performed using a PSI-BLAST algorithm [15] with Patt1 sequence [13] as an input (query). Default parameters were used, except for the "max target sequences" which was set to 1000. After eight iterations the set of resulting sequences showing statistically significant similarity was collected and used to construct the multiple sequence alignment (MSA) of Patt1 family. MSA was computed with the use of MUSCLE program [16, 17]. CDS (conserved domain search) [18] was used for assignment of the sequence to specific protein superfamily. Manipulation of sequence data was done with use of BioEdit [19] and SeaView [20] programs. ClustalX [21] was used to build phylogenetic trees (data not shown). TreeView [22] program was used to visualize them.

# Fold recognition analysis

Fold recognition methods are based on detection of compatibility of a given protein sequence with sequence profiles and/ or structures of proteins deposited in structural databases. In order to identify a suitable template for homology modeling of Patt1 structure the MSA of Patt1 family was submitted to GeneSilico Metaserver [23], which is a gateway to >30 third party methods of automated protein secondary structure prediction, solvation and disorder prediction (reference: https:// genesilico.pl/meta2/).

#### Patt1 structure modeling

The MODELLER program [24] available through GeneSilico Web Toolkit (https://genesilico.pl/toolkit/modeling) was used to construct an all-atom tertiary model of Patt1. Swiss-Pdb Viewer (SPDBV) program [25] was used for structure visualization and manipulation.

### Model quality assessment

All theoretical models must be thoroughly evaluated before using them as source of information of molecular architecture and composition of key structural elements. The state-of-theart methods of protein structure quality assessment specially designed to score theoretical models were used. Among them were MetaMQAPII [26] and ProQ [27] methods implemented in GeneSilico Web Toolkit (https://genesilico.pl/toolkit/).

The ProQ web server [28] (available at Stockholm Bioinformatics Center website: http://www.sbc.su.se/~bjornw/ ProQ/ProQ.html) was also used.

Energy minimization and molecular dynamics simulations

Molecular dynamics simulation and energy minimization were carried out using AMBER9 package [29]. Initial energy minimization in vacuo with unrestrained system was performed. Steepest descent algorithm was used for the first 6000 steps and conjugate gradient algorithm for following 24,000 steps. Cutoff value for treating long-range electrostatic interactions was 12.0. Simulation was performed with explicit solvent water model TIP3P. Counterions were added to neutralize net charge on protein surface. Long-range electrostatic interactions were evaluated using particle mesh Ewald summation method with periodic boundary conditions. Langevin thermostat method was used to control temperature of the system. The temperature was set to 300 K. The simulation duration was 12 ns, with time step (dt) of 2 fs.

## **Results and discussion**

## Patt1 sequence analysis

The sequence of Patt1 protein is 237 amino acid residues long. The set of sequences resulting from preliminary PSI-BLAST (default parameters, except for the "max target sequences = 1000" setting) search contained various members of GNAT family acetyltransferases from all domains of life. According to CDS results the region of similarity to GNAT superfamily was detected between approximately 112–187 amino acid residues. MSA was constructed and submitted to GeneSilico Metaserver. Disorder prediction indicated that the N-terminal (1–50 amino acid residues) and C-terminal regions (215–237 amino acid residues) of Patt1 protein were unstructured. The predicted regions of alpha helices and beta sheets conformed to the ( $\alpha$ )- $\beta$ - $\alpha$ - $\beta$ - $\beta$ - $\beta$ - $\alpha$ - $\beta$ - $\beta$ - $\beta$ - $\alpha$ - $\beta$ - $\beta$  pattern, which is in agreement with the canonical GNAT fold topology (Fig. 1). The prediction of first alpha helix is uncertain since it occurs

Patt1	DGLNVSIECKRVSGLEPATVDWAFDLTKTNMQTMYEQSEWG	KDREKREEMT-DDRAWYL	IAWENSSVPVAFSH
gi 262201662 ref YP_003272870.	MTAAEVIIDGLTYADIPRCAALEKQMGAEDSP	PPTAFRADIN-APYNTYF.	AARAE-PGGEVICYAC
gi 111023147 ref YP_706119.1	MTFRIEPMAAADAERCAELETLLEAGDGB	SAGAFRADIA-APHVHYT	VARDDTGHVVGYAG
gi 227541789 ref ZP_03971838.1	MSSSSVPSAAPSPATLEPLTAADVTRLEELEAQLEPGDSE	SAAAFAEDIR-SPWTYYV	GAE-RDGALIGYAC
gi 227505273 ref 2P_03935322.1	MKLEELTAADAPRCAELDIVLEPGESE	PAAABVQBIA-QPHTFYL	GVEDE-ETHTLVGMA
gi 19551827 ref NP_599829.1	MSEQFELMELRREDAGRCADLOQILBPGDND	PRDVBAVBFS-HPTNFMI	GAF-DEGYLVANAG
g1 119963453 ref YP_948593.1	MKLSPKLELAGVSLEDMTEADIPAWEALBRRLEPVDA	PLQMOFDOMA-QPETRRY	VAEVAGEIVAMA
g1 3/528082 ref NP_93142/.1	MNNISLLTPADLPSAFLVJKASHA - FP	SEKTEFGNOG-ERYLNYK	IAINEQLIGFAI
g1 1/0/68543 ref 2P_02902996.1		SERTEASNOG-ERYLNFQ	LTQNGKMAAFAI
g1 ///4//05 / FEI NP_030035.2		TRSIGRDCHQ-AGIPGWV	EEQAGQIIGHGV
gi 12290191421 rof170 04599569 1		CT POUL NOT PONDAUL	TC-EDEHTVENA
g1 2209310461 rof IVD 002507954		SKKENI PERODNKYSI VI	SCW-LDCPLVCVT
gi 12556580751ref12P 05403484 1		SPESIOWKIAA - NENTCYL	LAL-DGEOVICVA
gil238926143 ref ZP 04657903.1	MISERPLAPEDADAWACHERESIP-TP	SREDEWRICAL -NDFACHI	AAL-EGTAVICEAC
gil261878007 ref ZP 06004612.1	MIRFRALMPEDAEDWARVBYESEPTP	SREDEWREAS-NDFACY I	VAL-ADDSIIGFGG
gi 51894059 ref YP 076750.1	MGVPEIAVHPMTPADLDOVMEVPRLSYLTP	SREAFESELL-ORYTVYL	VAR-AGDRVVGFAC
gi 78044332 ref YP_359578.1	MELKDLPOVLDIEKLSYTNPWELKDLPOVLDIEKLSYTWE	SKASEMYEITENPLATYL	VAR-EGDKVIGYGG
gi 206900711 ref YP_002250683.	WQTQRKNSELSIEIRPMKFEDIDQVDEINKLSESNPW	SRESSERDISSNRIAHYF	VAT-YENKVVGFVG
gi 217967351 ref YP_002352857.	WQDQKESNELSVEIRPMKFEDIDQVDEINKLSFSNPW	SRESPERELSSNRIAHYF	VAI-HENKIIGFVC
gi 23098102 ref NP_691568.1	MNNITMEPMELADIPEVLHVBNACGQTP	TTDIFYQELIENAHAYYY	VVE-VERTVVGYIC
gi 229555411 ref ZP_04443200.1	MNESGKVVLENAELEDLPGILRVEKAAFTSP	TETAFRNELVLNEFAAYI	VIE-LDAVIIGYAG
gi 46908312 ref YP_014701.1	MSLDEALLFREATVTDLKSIMNVDNAAFTVB	TEAAFRNEFIINQYAYYL	LAI-YKEQVVGYAG
gi 21283703 ref NP_646791.1	MDQQSKEQLNINEWTKEDVPOWFDNBRRSBN-DSS	TIDAFYHBIEQNNFAKYF	VLE-FEQQIIGYLC
g1 27468569 ref NP_765206.1	WVKPTREQLNINKUSVEDVPKWFDHORNSDS-HSS	SIDAFYHDIENNEFATYF	VIE-FSDKIIGYVe
g1 239637783 ref 2P_04678747.1	MRPTEEQLNIRK/SKDVPAWFDIDRSSSN-DSS	SIDAEYHBIEKNEFANYF	VIE-FNHTIIGALG
g1 261409085 ref YP_003245326.		TEEASHNOIKLNHFARM	IME-YEGRTVGHAG
g1 16/462251 Te1 2P_0232/340.1		TO BE AN AN A TO A TO A TO A TO A TO A TO A T	WME-VHGDIAGHGG
gi 226310101   rof VP 003769995		DUDADUNG T DNDNADYU	WW A-HONDITANCO
gi 1220310101 [Tel   12_002703995.		T D S D Y D U U P NO Y NUYU	LAIDENCETTEEC
gi 11637630121ref17P 02170075 1		SPREANETVENOFAVAT	MA
gil121535447   ref   ZP 01667257.1	MSRRKWVISPMTADDIDAVLEVBRASDT-TP	SRAMBAABVADNDLAYNL	GE-ADGAVVGYAG
gi 157691317 ref YP 001485779.		KKESEVYETLHNOYSHYF	LIE-EDEOPVCYCC
gi 52079049 ref YP 077840.1	MNTOMVIRDMNPKDLDOWFEIPKSSFSSP	KKESCHHOLFHNMYAHYL	VLE-VEEOVVGYC
gil1546850931reflyp 001420254	MKTKTAIRSMRPEDIDCIYEIETASETSP	TKDSEYHELLENPYAHYL	VIE-KDGCLACYCC
gi 15613110 ref NP_241413.1	WAEKKERIRFMTVDDVDAWVRVBEQTETVPWAEKKERIRFMTVDDVDAWVRVBEQTETVPWAEKKERIRFMTVDDVDAWVRVBEQTETVPWAEKKERIRFMTVDDVDAWVRVBEQTETVPWAEKKERIRFMTVDDVDAWVRVBEQTETVPWAEKKERIRFMTVDDVDAWVRVBEQTETVPWAEKKERIRFMTVDDVDAWVRVBEQTETVPWAEKKERIRFMTVDDVDAWVRVBEQTETVPWAEKKERIRFMTVDDVDAWVRVBEQTETVPWAEKKERIRFMTVDDVDAWVRVBEQTETVPWAEKKERIRFMTVDDVDAWVRVBEQTETVPW	SREAFINEVINNOFARYV	VYE-VGEQIVGYCC
gi 15613110 ref NP_241413.1  gi 205372356 ref ZP_03225170.1	WARSREIEGVEVHIADQEIMCREMQLADLDEVLVVENDAFTLPWARSREIEGVEVHIADQEIMCREMQLADLDEVLVVENDAFTLPWARSREIEGVEVHIADQEIMCREMQLADLDEVLVVENDAFTLPWARSREIEGVEVHIADQEIMCREMQLADLDEVLVVENDAFTLPWARSREIEGVEVHIADQEIMCREMQLADLDEVLVVENDAFTLPWARSREIEGVEVHIADQEIMCREMQLADLDEVLVVENDAFTLPWARSREIEGVEVHIADQEIMCREMQLADLDEVLVVENDAFTLPWARSREIEGVEVHIADQEIMCREMQLADLDEVLVVENDAFTLPWARSREIEGVEVHIADQEIMCREMQLADLDEVLVVENDAFTLPWARSREIEGVEVHIADQEIMCREMQLADLDEVLVVENDAFTLP	SREAFINEVTNNOFARYV KRSAFENELEHNSFARYR	VYE-VGEQIVGYCG VLV-EKEQIIGYCG
gi 15613110 ref NP_241413.1  gi 205372356 ref ZP_03225170.1 gi 163938250 ref YP_001643134.	MARKERRIGPUTVDDVDAUNUVEQDT-VD	SREAFINEVTNNQFARYV KRSAFENELEHNSFARYR TADAFHRELEVNEHAHYV	VYE-VGEQIVGYCG VLV-EKEQIIGYCG VLE-KDGLVIGYCG
gi 15613110 ref NP_241413.1  gi 205372356 ref ZP_03225170.1 gi 163938250 ref YP_001643134. gi 165873317 ref ZP_02217921.1		SREAFINEVTNNOFARYV KRSAFENELEHNSFARYR TADAFHRELEVNEHAHYV TADAFHRELTMNEHAHYV	VYE-VGEQIVGYCG VLV-EKEQIIGYCG VLE-KDGLVIGYCG VLE-KDGRVIGYCG
gi 15613110 ref NP_241413.1  gi 205372356 ref ZP_03225170.1 gi 163938250 ref YP_001643134. gi 165873317 ref ZP_02217921.1	MARKERREFERDERVODVDAUNTVECTET-VE 	SREAFINEVTNNOFARYV KRSAFENELEHNSFARYR TADAFHRELEVNEHAHYV TADAFHRELTMNEHAHYV	VYE-VGEQIVGYCG VLV-EKEQIIGYCG VLE-KDGLVIGYCG VLE-KDGRVIGYCG
gi  15613110 ref NP_241413.1 gi  205373356 ref PP_03225170.1 gi  163938250 ref YP_001643134. gi  165873317 ref YP_00217921.1 Patt1		SREAFINEVTNNQFARYV KRSAFENELEHNSFARYR TADAFHRELEVNEHAHYV TADAFHRELTMNEHAHYV FREALQFEIDDSSPSM	VYE-VGEQIVGYCG VLV-EKEQIIGYCG VLE-KDGRVIGYCG VLE-KDGRVIGYCG SGCCGEDCSYEI
ği   15613110   rof   NP 241413.1     gi   205372356   ref   ZP 03225170.1     gi   163938250   ref   XP 001643134.     gi   16587317   ref   ZP 02217921.1     Patt1   Patt1	MAEKKERIGFURVOD VAUVENUS VUOD VUOD VUOD VUOD VUOD VUOD VUOD VUOD	SREAFINEVTNNOFARNY KRSAFENETEHNSFARM TADAFHREIEVNEHAHYV TADAFHREITMNEHAHYV FREAIQSEIDDSSPSM MSR-HOITISOFRHM	VYE-VGEQIVGYCG VLV-EKEQIIGYCG VL-E-KDGLVIGYCG VL-E-KDGRVIGYCG SGCCGEDCS-YEI YQPSGADAF-TMM
gi  15613110  ref  NP 241413 1 gi  205372356  ref  ZP 03225170.1 gi  205372356  ref  ZP 01643134. gi  163938250  ref  ZP 001643134. gi  165873317  ref  ZP 02217921.1 Patt1 gi  262201662  ref  YP 003272870. gi  21023147  ref  YP 706119.1  gi  22741780  ref  YP 706119.1  gi  22841780  ref  YP 706119.1  gi  23841780  ref  YP 706119.1  gi  XP 70841780  ref  YP 70841	MARSKEIRGEWTUD UD AUR WEIGENT - VC 	SREAFINEVTNNOFARYV KRSAFENEJEHNSFARYR TADAFHREJEVNEHAHYV FREALOF-EIDDSSPSM MSRHGE-TTSGERRHY YRR-EGTTSGERRHY YRR-EGVOIGNEYN	VYE-VGEQIVGYC VLE-KDGLVIGYC VLE-KDGLVIGYC SGCCGEDCS-YEI YQPSGADAF-TMM YQPSGADAF-TMM
gi   15613110   ref   NP _ 241413.1 gi   205373356   ref   PP _ 03225170.1 gi   163938250   ref   PP _ 01643134. gi   165873317   ref   ZP _ 02217921.1 Patt1 gi   262201662   ref   YP _ 003272870. gi   111023147   ref   YP _ 706119.1 gi   227541789   ref   ZP _ 03971838.1 gi   22756273   ref   ZP _ 039735322 1	MAEKKERIGPUTVDDVDAUVRVEQETT-VD- 	SREAFINEVTNOPARVY KRSAFEMENENSFRMR TADAFREJEVSERAF TADAFREJEVSERAF SREALOFEIDDSSPSM VSR-HGETISGTRHY VRR-EGEEIVGTRKY VRR-EGEEIVGTRKY VRE-HGEVIGVRNN	Y E - VGEQIVGY VL KEQIVGY VL E - KDGLVIGY VL E - KDGLVIGY SG CCG⊡CS YEI YQ PSGADAF - TMR YQ PSGADAF - TMR YQ PSGADAF - TMR YQ PSGADAF - TMR YQ PSGADAF - TMR
<pre>gi   15613110  ref   NP 241413.1] gi   205372356  ref   ZP_03225170.1 gi   163938250  ref   ZP_03225170.1 gi   163938250  ref   ZP_001643134. gi   165873317  ref   ZP_02217921.1 Patt1 gi   262201662  ref   ZP_003272870. gi   217241789  ref   ZP_03971838.1 gi   227505273  ref   ZP_03955322.1 gi   197505273  ref   ZP_03955322.1 gi   19751827  ref   ZP_03955322.1</pre>	MARSREIEGVEVHGUDADAUVRVESOTT-VE	SREATING TANGO PARTY KRSATENETEHSSANN TADATHATLEVNEHAMY TADATHATLEVNEHAMY FRALQT-EIDSSPSM YSR-HGT-TTSGTRAHY YRR-EGT-TTSGTRAHY YRG-HGT-VOLGYRSN YEA-YGT-IKTGTRAH YEA-YGT-IKTGTRAH	Y E - VGEQIVGYC VL E + EQIIGYC VL E + EQIIGYC VL E + DGRVIGYC SG CCGEDCS YEI YO PSGADAF - TM YO PSGADAF - TM YO PSGADAF - TM YO PSGADAF - TM YO PSGADAF - TM
gi  15613110 ref NP_241413.1 gi  205372356 ref PP_03225170.1 gi  163938250 ref PP_03225170.1 gi  163938250 ref YP_001643134. gi  262201662 ref YP_003272870. gi  11023147 ref YP_706119.1 gi  227541789 ref PP_03971838.1 gi  227505273 ref YP_0353522.1 gi  15551827 ref NP_599829.1  gi  15551827 ref YP_68593.1		SREAFINGUTNOPARU KRSAJENDJENSFARM TADATHEJEVNEHAMU TADATHEJEVNEHAMU FREALQI-EIDSSFSM VSR-HGI-EIDSSFSM VSR-HGI-EIVGURN VRR-COI-EIVGURN VRR-COI-EITGURN VRR-FGI-KILAVRN VRA-FGI-KILAVRN	Y E - ∨ GEO IVGYC VL E + CDGIVGYC VL E + KDGIVIGYC VL - E - KDGIVIGYC VL B - KDGIVIGYC SG CCG⊡CS YEI VC PSGADAF TW YC PSGADAF - TW YC PSGADAH - TW YC PSGADAH - TW YR PSGADAH - TW
<pre>i   15613110  rof NP 241413 1 gi   205372356  ref   ZP 03225170.1 gi   163938250  ref   ZP 03225170.1 gi   163873317  ref   ZP 032217921.1 Patt1 gi   262201662  ref   YP 003272870. gi   217241789  ref   ZP 03971838.1 gi   227505273  ref   ZP 03971838.1 gi   1951827  ref   ZP 03935322.1 gi   1951827  ref   ZP 03935322.1 gi   19563827  ref   ZP 03935323.1 gi   37528082  ref   NP 948593.1  gi   37528082  ref   NP 94827.1</pre>		SREAT IN JUTN OF AN UN KROATEN JEH SFAMTR TADATHELEV EHANN TADATHELEV EHANN FRAIGH - EIDSSPS VSR - HGE - TISGIRHM VRR - GE - EIDSSPS VRR - GE - VOLGVRN VRA - FGE - KILAVRN VRA - FGE - KILAVRN VRA - FGE - CIHVRRN VRA - FGE - ONOVSVRH	$ \begin{array}{c} \mathbf{y} = -\mathbf{y} = -\mathbf{y} \in \mathbf{g} \in \mathbf{y} \in$
gi  15613110 ref NP_2414131] gi  205372356 ref ZP_03225170.1 gi  163938250 ref ZP_03225170.1 gi  163938250 ref ZP_02217921.1 Patt1 gi  262201662 ref YP_003272870. gi  21023147 ref YP_706119.1  gi  227541789 ref ZP_03971838.1 gi  227505273 ref ZP_03953222.1] gi  139551827 ref YP_94593.1  gi  31528082 ref YP_931427.1  gi  37528082 ref XP_931427.1  gi  37508543 ref ZP_03029961		SREAT IN GTAN OF AR Y KRSAFENEJEHNSFARN TADAFHELIEVNEHAMY TADAFHELITNEHAMY SRHGF-TTSGTRHY YRR-GF-TTSGTRHY YRR-GGT-TTSGTRHY YRR-GGT-KIAVRAN YEA-YGJ-KIAVRAN YEA-YGJ-KIAVRAN YEA-GG-NOVSVRHY YEC-GG-NEATTRNY	VY E-VGEQ YV C VI V- EKEQ I ( Y C VI E - KDGI V ( G Y C VI E - KDGI V ( G Y C VI E - KDGI V ( G Y C V) PSGAD AF T V R DSGAD AF T V R DSGAD AF T V T T G KED AF T A
gi   15613110   i r f NP 241413     gi   205372356   i r f IZP 03225170.1     gi   163938250   i r f IZP 03225170.1     gi   165873317   i r f IZP 032217921.1     Patt1   gi   262201662   i r f IZP 03272870.     gi   1262201662   i r f IZP 03971838.1   gi     gi   2271791   i r f IZP 03971838.1   gi     gi   2275172.1   i r f IZP 03971838.1   gi     gi   2275172.3   i r f IZP 03971838.1   gi     gi   1951827   i r f IZP 03971838.1   gi     gi   19505273   i r f IZP 048593.1   gi     gi   1950453   i r f IZP 048593.1   gi     gi   170768543   i r f IZP 02902996.1   gi     gi   17077765   i r f IZP 048535.2   i		SPERTINGTRNOPARUN RESTENSTEHSSAMMR TADAFHREIEVSEAMMR TADAFHREIEVSEAMMR PREATOF - EIDDSSS WSR-HGE-TTSGTRHY WRE-GE-EIGTRHY WRE-GE-VOLGVRRNY VEA-FGE-VOLGVRRNY VEA-FGE-CITHVRRNY VEA-FGE-CITHVRRNY VEA-FGE-CITHVRRNY VEA-FGE-CITHVRRNY VEA-FGE-CITHVRRNY VEA-FGE-NGVSVRRHY VEC-AGE-NEIGRRNY	Y
ği   15613110   ref NP 241443 1 gi   205372356   ref   ZP 03225170.1 gi   263932250   ref   ZP 03225170.1 gi   163932250   ref   ZP 001643134. gi   165873317   ref   ZP 002217921.1 Patt1 gi   262201662   ref   YP 003272870. gi   217541789   ref   ZP 03971838.1 gi   227505273   ref   ZP 0395322.1 gi   19953453   ref   ZP 0395323.1 gi   19551827   ref   ZP 039533.1 gi   19563453   ref   ZP 931427.1 gi   17768543   ref   ZP 03029963.2 gi   177747765   ref   ZP 03029963.2 gi   177747765   ref   ZP 03029963.2 gi   15298429   ref   ZP 03029963.2 gi   177747765   ref   ZP 03029963.2 gi   15298429   ref   ZP 030350715.2 gi   15298429   ref   ZP 03350715.2 gi   15298429   ref   ZP 0350715.2 gi   15298429   ref   ZP 03350715.2 gi   15298429   ref   ZP 03350715.2 gi   15298429   ref   ZP 03350715.2 gi   15298429   ref   ZP 03350715.2 J 15208429   ref   ZP 03350715.2 J 15308429   ref   ZP 03350715.2 J 15308429   ref   ZP 03350715.2 J 15308429   ref   ZP 0356715.2 J 1530845715.2 J 1530845715.2 J 1530845715.2 J 1530845715.2 J 1530845715.2 J 1530845715.2 J 153		SREATINGTING PARTY KRSATENTIEHNSFARMY KRSATENTIEHNSFARMY TADATHEITNIEHAMY FRAIQT-TISGTRHY YRR-EGT-TISGTRHY YRR-EGT-TISGTRHY YRR-EGT-VOLGVRNY YEA-FGT-VOLGVRNY YEA-FGT-EQIHVRDY YEA-FGT-EQIHVRDY YEA-FGT-EQIHVRDY YEA-FGT-NEIGTRNY YES-LGT-NEIGRRAY YHS-EGT-NEIGRRAY	Y E-VGEQ IVA V E-KQIIQYC V E-KDGVIGYC V E-KDGVIGYC V E-KDGVIGYC Y PSGADAPTW Y PSGADAPTW Y PSGADAPTW Y PSGADAPTW Y PSGADAYTW Y PSGADAYTW Y PSGADAYTW Y PSGADAYTW Y PSGADAYTW Y PSGADAYTW Y PSGADAYTW Y PSGADAYTW Y PSGADAYTW
iiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiii		SPERTINGTNO OFAR W RESPENSION OFAR W TADATHRIPY SHAN W TADATHRIPY SHAN W TADATHRIPY SHAN W TADATHRIPY SHAN W SR HGT TISGIRHY YR GI EIDDSS SM YR FGI VOLGVRNN WEA-FGI VOLGVRNN WEA-FGI KILAVRNN YEA-FGI KILAVRNN YEA-FGI NIGRRNY YES-GI NEIGRRNY YES-GI NEIGRRNY YES-GI NEIGRRNY YES-GI NEIGRRNY YES-GI NEIGRRNY YES-GI NEIGRRNY YES-GI NEIGRRNY YES-GI NEIGRRNY	YU E-VECQUIGYC       YL E-KDGLVIGYC       YL PSGADAPT       YL YL YL T       YL
gi   15613110   ref NP 241413 1     gi   205372356   ref   ZP 03225170.1     gi   26393250   ref   ZP 03225170.1     gi   16393250   ref   ZP 03225170.1     gi   16587317   ref   ZP 032217921.1     Patt1     gi   262201662   ref   YP 003272870.     gi   262201662   ref   YP 003272870.1     gi   227541789   ref   ZP 03971838.1     gi   227505273   ref   ZP 03971838.1     gi   19963453   ref   XP 948593.1     gi   170768543   ref   ZP 02902906.1     gi   17774765   ref   NP 636035.2     gi   238019142   ref   ZP 001350715.3     gi   233019142   ref   ZP 0202906.1     gi   123031046   ref   ZP 04599566.1     gi   233019142   ref   ZP 04599568.1		SREAT IN TYTN O PAR Y KRSAFENDIEH SFAR Y KRSAFENDIEH SFAR Y TADAFHRIEV MEHAHY YRE-EGT-TISGIRHY YRE-EGT-TISGIRHY YRE-EGT-VOLGVRNY YEA-FGT-VOLGVRNY YEA-FGT-NEIGIRHY YES-IGT-NEATIRNY YES-IGT-NEATIRNY YES-IGT-NEATIRNY YES-IGT-NEATIRNY YES-IGT-NEATIRNY YES-IGT-NEATIRNY YES-IGT-NEATIRNY YES-IGT-NEATIRNY YES-IGT-NEATIRNY YES-IGT-NEATIRNY YES-IGT-NEATIRNY	VY
gi   15613110   ref  NP_2414131     gi   205372356   ref  ZP_03225170.1     gi   163938250   ref  ZP_03225170.1     gi   163938250   ref  ZP_03225170.1     gi   163938250   ref  ZP_03225170.1     gi   165873317   ref  ZP_03217921.1     Patt1   gi   22201662     gi   22201662   ref  YP_03183.1     gi   227505273   ref  ZP_03935322.1     gi   127505273   ref  ZP_03935322.1     gi   137528082   ref  YP_948593.1     gi   137528082   ref  ZP_03022996.1     gi   170768543   ref  ZP_03022996.2     gi   170768543   ref  ZP_0302596.2     gi   1238019142   ref  ZP_045959754.     gi   120931046   ref  ZP_00507954.     gi   12093046   ref  ZP_00350755		SREAT IN GTAN OF AN Y KRSAFENEJEN SFARNY KRSAFENEJEN SFARNY TADATHELITNEHAMY TRAATAS THE TAME HAMY YREAGT EIDSSFSM YREAGT TIGGTRHY YREAGT TIGGTRHY YREAGT TVGGTRHY YEA-YGG KIAVRAN YEA-YGG KIAVRAN YEA-YGG NOUSYRAN YEG-LGG NEATRAN YHS-GG NEIGRRAD YER-YGG NEIGRRAD YER-YGG NEIGRRAD YER-YGG NEIGRRAD YER-JGG TVKGTRAD YER-JGG TVKGTRAD	$ \begin{array}{l} \mathbf{y} = -\mathbf{y} = -\mathbf{y} \in \mathbf{g} \in \mathbf{y} \in \mathbf{y} \in \mathbf{y} \\ \mathbf{y} = -\mathbf{y} = \mathbf{k} \in \mathbf{g} \in \mathbf{y} \in \mathbf{y} \in \mathbf{y} \in \mathbf{y} \\ \mathbf{y} = -\mathbf{k} = \mathbf{k} \in \mathbf{g} \in \mathbf{y} \in \mathbf{y} \in \mathbf{y} \in \mathbf{y} \\ \mathbf{y} = -\mathbf{k} = \mathbf{k} \in \mathbf{g} \in \mathbf{y} \in \mathbf{y} \in \mathbf{y} \\ \mathbf{y} = -\mathbf{k} \in \mathbf{g} \in \mathbf{k} \in \mathbf{y} \in \mathbf{y} \in \mathbf{y} \\ \mathbf{y} = -\mathbf{k} \in \mathbf{g} \in \mathbf{k} \in \mathbf{k} \\ \mathbf{y} = -\mathbf{k} \in \mathbf{g} \in \mathbf{k} \in \mathbf{k} \\ \mathbf{y} = -\mathbf{k} \in \mathbf{g} \in \mathbf{k} \in \mathbf{k} \\ \mathbf{y} = -\mathbf{k} \in \mathbf{g} \in \mathbf{k} \\ \mathbf{y} = -\mathbf{k} \in \mathbf{g} \in \mathbf{k} \in \mathbf{k} \\ \mathbf{y} = -\mathbf{k} \in \mathbf{g} \in \mathbf{k} \\ \mathbf{y} = -\mathbf{k} \in \mathbf{k} = -\mathbf{k} \\ \mathbf{y} = -\mathbf{k} \in \mathbf{k} = -\mathbf{k} \\ \mathbf{y} = -\mathbf{k} = \mathbf{k} = -\mathbf{k} \\ \mathbf{y} = -\mathbf{k} = -\mathbf{k} \\ \mathbf{x} = -\mathbf{k} = -\mathbf{k} \\ \mathbf{x} = -\mathbf{k} \\ $
gi   15613110   rof NP 241413 1     gi   205372356   ref   ZP 03225170.1     gi   205372356   ref   ZP 03225170.1     gi   163938250   ref   ZP 03225170.1     gi   165873317   ref   ZP 03217921.1     Patt1     gi   262201662   ref   YP 001643134.1     gi   262201662   ref   YP 003272870.1     gi   262201662   ref   YP 003272870.1     gi   227501273   ref   ZP 03971838.1     gi   227505273   ref   ZP 03971838.1     gi   195582801   ref   XP 0395829.1     gi   19558082   ref   XP 048593.1     gi   170768543   ref   ZP 048593.1     gi   170768543   ref   ZP 048593.1     gi   2350802   ref   XP 001350715.     gi   2350519142   ref   XP 001350715.     gi   23505075   ref   XP 00357954.3     gi   23563075   ref   ZP 0403484.1     gi   23563075   ref   ZP 0567903.4		SREAT IN TYTN OPAN WY REATENTEN TEN OPAN WY REATENTEN TEN SEAN R TADATHELEYNEHAM WY TADATHELEYNEHAM WY FRAIGT-TISGIRHM YRR-EGT-TISGIRHM YRR-EGT-VOLGVRNN YRR-FGT-VOLGVRNN YEA-FGT-FTIATRNN YLR-FGT-NIIGRRDW YES-EGT-NEIGRRNY YES-EGT-NIIGRRDW YES-EGT-NIIGRRNY YES-EGT-NIIGRRDW YES-EGT-TYRGRRNY YES-EGT-TYRGRRNY YES-EGT-TYRGRRNY YES-EGT-TYRGRRNY YES-EGT-TYRGRRNY YES-EGT-TYRGRRNY YES-EGT-TYRGRRNY YES-EGT-TYRGRRNY	Y Y = V Y + C V
<pre>gi  15613110 ref  NP 241413 1 gi  205372356 ref  ZP 03225170.1 gi  205372356 ref  ZP 03225170.1 gi  263932250 ref  ZP 001643134. gi  2652201662 ref  YP 003272870. gi  262201662 ref  YP 003272870. gi  227541789 ref  ZP 03971838.1 gi  227505273 ref  ZP 03971838.1 gi  1551827 ref  ZP 0395322.1 gi  15505273 ref  ZP 03955322.1 gi  1551827 ref  ZP 0395532.1 gi  37528082 ref  ZP 0302598.1 gi  37528082 ref  ZP 03025986.1 gi  37747765 ref  ZP 04599568.1 gi  23019142 ref  ZP 04599568.1 gi  225658075 ref  ZP 04595703.1 gi  23021046 ref  ZP 0457903.1 gi  2256143 ref  ZP 04657903.1 gi  2261878007 ref  ZP 0600661.1</pre>		SREAT IN GYTN O PAR Y KRSAFENEIEN SFARNY KRSAFENEIEN SFARNY TADAFHEITNEHAMY FRAIGENTSGIRHY YRE-GG-EIGHAY YRE-GG-VOIGVRAY YRE-GG-VOIGVRAY YEA-YGF-KIAVRAY YEA-YGF-KIAVRAY YEA-GG-NEATRAY YEC-GG-NEATRAY YES-LGG-NEATRAY YHS-GG-NEATRAY YHS-GG-NEATRAY YHS-GG-NEATRAY YHS-GG-NEATRAY YHS-GG-NEATRAY YHS-GG-VEIGRAPY YRH-GG-THIGCIYG YRH-GG-FRAGVRAY	$ \begin{array}{c} v_{T} = -v_{S} = v_{S} \in v_{S} \in v_{S} v_{S} \\ v_{L} = -v = K \in V \in V \in V \in V \\ v_{L} = -k = K \in V \in V \in V \in V \\ v_{L} = -k = K \in V \in V \in V \in V \\ v_{L} = -k = K \in V \in V \in V \\ v_{L} = -k \in V \in V \in V \\ v_{L} = -k \\ v_{L} $
gi   15613110   rof NP 241413   1     gi   205372356   ref   ZP 03225170.1     gi   163938250   ref   YP 001643134.     gi   165873317   ref   ZP 032217921.1     Patt1     gi   262201662   ref   YP 001643134.     gi   262201662   ref   YP 001643134.     gi   262201662   ref   YP 003272870.     gi   262201662   ref   YP 003272870.     gi   26741789   ref   ZP 03971838.1     gi   227505273   ref   ZP 03971838.1     gi   3528082   ref   XP 948593.1     gi   37528082   ref   XP 948593.1     gi   170768543   ref   ZP 03035.2     gi   170768543   ref   XP 039584.     gi   2399249   ref   XP 048593.1     gi   2399429   ref   XP 04359568.1     gi   220931046   ref   XP 0459568.1     gi   228926143   ref   ZP 04657903.1     gi   238926143   ref   ZP 04657903.1     gi   23894039   ref   ZP 04657903.1     gi   23894039   ref   ZP 04657903.1     gi   23894039   ref   ZP 076050.1		SREAT IN TYTM OPAN WY REATER SAME AND	$ \begin{array}{c} \mathbf{y} = - \mathbf{y} = - \mathbf{y} \in \mathbf{y}$
<pre>gi   15613110   ref NP 241413 1 gi   205372356   ref   ZP 03225170.1 gi   163938250   ref   ZP 03225170.1 gi   163938250   ref   ZP 032217921.1 Patt1 gi   262201662   ref   YP 003272870. gi   227541789   ref   ZP 03971838.1 gi   227505273   ref   ZP 03971838.1 gi   2351827   ref   ZP 03971838.1 gi   19963453   ref   ZP 03971838.1 gi   19953453   ref   ZP 048593.1 gi   137528082   ref   ZP 048593.1 gi   13728042   ref   ZP 045958.1 gi   15298429   ref   ZP 0459568.1 gi   22031046   ref   ZP 04557958.1 gi   220301046   ref   ZP 04557958.1 gi   2238015142   ref   ZP 0457903.1 gi   2380459   ref   ZP 0457903.1 gi   2561878007   ref   ZP 0604612.1 gi   51894059   ref   ZP 060750.1 gi   58044332   ref   YP 0750.1 gi   5804059   ref   YP 0750.1 gi   208071   ref   YP 0750.1 gi   208071   ref   YP 0750.1</pre>	MARSREIEGVEVHIAQUEINGEWQLADIDELUVUWINDAT	SREAT IN TYTN OPAN Y KRSAFENDIENSFAN Y KRSAFENDIENSFAN Y TADATHEIEWSEHAMY TADATHEITNEHAMY FRALOF EIDSSPSM YSR-HGE TTSGTRHY YRR-GG EIVGTRKY YRG-HGE VOLGYRN YEA-FGF VOLGYRN YEA-FGF KTLAVRN YEA-FGF KTLAVRN YEA-FGF NEIGRRPY YES-LGG NEIGRRPY YHS-GG NEIGRRPY YHS-GG NEIGRRPY YRH-GG NEIGRRPY YRH-GG NEIGRRPY YRH-GG NEIGRRPY YRH-GG NEIGRRY YRH-GG NEIGRRY YRH-GG NEIGRRY YRH-GG SAVGYRKY YEG-LGG SAVGYRKY	Y
<pre>i   15613110   rof NP 241413 1   gi   205372356   ref   ZP 03225170. 1 gi   205372356   ref   ZP 03225170. 1 gi   163938250   ref   YP 001643134. gi   165873317   ref   ZP 02217921. 1 Patt1 gi   26201662   ref   YP 003272870. gi   2127541789   ref   ZP 03971838. 1 gi   227505273   ref   ZP 03971838. 1 gi   2558082   ref   ZP 03971838. 1 gi   3528082   ref   ZP 03971838. 1 gi   37528082   ref   ZP 048593. 1 gi   229931046   ref   ZP 04959568. 1 gi   2289429   ref   ZP 04599584. 1 gi   22586075   ref   ZP 04657903. 1 gi   2588707   ref   ZP 0457903. 1 gi   25896143   ref   ZP 0457903. 1 gi   25187807   ref   ZP 045578. 1 gi   206900711   ref   YP 0025578. 1 gi   2067351   ref   YP 0025578. 1 gi   2075751   ref   YP 076750. 1 gi   2075751   ref   YP 076750. 1 gi   2075751   ref   YP 076750. 1 gi   2075751   ref   YP 07675578. 1 gi   2075751   ref   YP 057578. 1 gi   2075751   ref   YP 057578. 1 gi   207575751   ref   YP 057578. 1 gi   207575751   ref   YP 057578. 1 gi   207575751   ref   YP 057578. 1 gi   20757575751   ref   YP 057575757575757575757575757575757575757</pre>	MARSREIGGVHIAUQELRÜGEVGAUDEVLÖVENUN UND VALUEVAUSANAAIAN 	SPERTINGTNO OFAR W REATEN SPANNE TADAFHREIEVSEAM R TADAFHREIEVSEAM FRAIOF - EIDSSPSM SRAHGE-TTSGTRHY RRAEGE-TSGTRHY RRAEGE-TSGTRHY RRAFGE-VOLGVRKY VRAFGE-VOLGVRKY VRAFGE-NCUSYRRHY VEA-FGE-SUGVRRY VEA-FGE-SUGVRRY VEA-FGE-SUGVRRY VEA-FGE-SUGVRRY VEA-FGE-SUGVRRY VEA-FGE-SUGVRRY VEA-FGE-SUGVRRY VEA-FGE-SUGVRRY VEA-FGE-SUGVRRY VEA-FGE-SUGVRRY VEA-FGE-SUGVRRY VEA-FGE-SUGVRRY VEA-FGE-SUGVRRY VEA-FGE-SUGVRRY VEA-FGE-SUGVRRY VEA-FGE-SUGVRRY VEA-FGE-SUGVRRY VEA-FGE-SUGVRRY VEA-FGE-SUGVRRY VEA-FGE-VEIGRAN VEA-FGE-VEA VEA-FGE-VEA VEA-FGE-VEA VEA-FGE-VEA VEA VEA-FGE-VEA VEA VEA VEA VEA VEA VEA VEA	Y
<pre>i i 5613110 r e r NP 241413 1 gi 205372356 r e f   ZP 03225170.1 gi 163932250 r e f   ZP 03225170.1 gi 163932250 r e f   ZP 03225170.1 gi 165873317 r e f   ZP 032217921.1 Patt1 gi 225201662   r e f   YP 003272870. gi 27541789 r e f   ZP 03971838.1 gi 227505273   r e f   ZP 03971838.1 gi 12551827 r e f   ZP 03971838.1 gi 19551827 r e f   ZP 0395322.1 gi 170768543 r e f   ZP 048593.1 gi 170768543 r e f   ZP 048593.1 gi 177765 r e f   ZP 048593.1 gi 177765 r e f   ZP 04350715. gi 122989429 r e f   ZP 04399568.1 gi 125989429 r e f   ZP 04399568.1 gi 238019142 r e f   ZP 0457903.1 gi 238019142 r e f   ZP 0457903.1 gi 25658077 r e f   ZP 0604612.1 gi 151894059 r e f   ZP 0604612.1 gi 151894059 r e f   YP 07575.1 gi 22080711 r e f   YP 0755.1 gi 22050711 r e f   YP 022550685.1 gi 22090711 r e f   YP 002352857. gi 237973 r e f   ZP 0255687.1 gi 2205011 r e f   YP 002352857. gi 237973 r e f   ZP 0575.1 gi 23809102 r e f   YP 07578.1</pre>	MARSREIE GVEVHIADQ EINGEN QIAD DUG LIVVI DUDA VRÜEQTE - VD 	REBATING     STATUS       RESATENDIENS     SPANN       KRSAFENDIENS     SPANN       KRSAFENDIENS     SPANN       KRSAFENDIENS     SPANN       TADATHRIE     SPANN       TRESE     SPANN       TRESE     SPANN       TEG	Y =
gi   15613110   rof NP 241413   1     gi   205372356   ref   ZP 03225170.1     gi   163938250   ref   YP 001643134.     gi   16587317   ref   ZP 03225170.1     Patt1     gi   262201662   ref   YP 003272870.     gi   2275173   ref   ZP 03371838.1     gi   2275173   ref   ZP 03371838.1     gi   227541789   ref   ZP 03371838.1     gi   22754378   ref   ZP 03371838.1     gi   22754378   ref   ZP 03373322.1     gi   139528082   ref   NP 948593.1     gi   37528082   ref   NP 948593.1     gi   375289429   ref   NP 0035075.5     gi   226931046   ref   SP 06035.2     gi   226931044   ref   SP 00457903.1     gi   22695103   ref   ZP 04657903.1     gi   22699104   ref   SP 045578.1     gi   78044322   ref   YP 076750.1     gi   206900711   ref   YP 0025578.1     gi   23098102   ref   YP 035578.1     gi   23098102   ref   NP 691568.1     gi   23098102   ref   NP 691588.1    gi   23098102   ref   NP 691588.1 <td>MARSREIGGVHIAQUERINGEWGLURADUR VUD VANDAVRIECT - VD</td> <td>S R E R I M VT M O PAR W R E R T M FRIENS FAM WR TADAFHREIEV PAR M TADAFHREIEV PAR M TADAFHREIEV E HAN V TADAFHREIEV S HAN V FRANCE S R HGF TISGIRHY WR - FG TISGIRHY WR - FG VOLGVRKN WR - FG VOLGVRKN WE - FG RIGRNN YE - FG RIGRNN YE RIGRNN YE RIGRNN WE FG NIGRNN YE RIGRNN YE RIGRNN YE SIGRNN YE SIGRNN Y</td> <td>Y</td>	MARSREIGGVHIAQUERINGEWGLURADUR VUD VANDAVRIECT - VD	S R E R I M VT M O PAR W R E R T M FRIENS FAM WR TADAFHREIEV PAR M TADAFHREIEV PAR M TADAFHREIEV E HAN V TADAFHREIEV S HAN V FRANCE S R HGF TISGIRHY WR - FG TISGIRHY WR - FG VOLGVRKN WR - FG VOLGVRKN WE - FG RIGRNN YE - FG RIGRNN YE RIGRNN YE RIGRNN WE FG NIGRNN YE RIGRNN YE RIGRNN YE SIGRNN YE SIGRNN Y	Y
<pre>i i 5613110   rof NP 241413 1 gi 205372356   ref   ZP 03225170.1 gi 163938250   ref   ZP 03225170.1 gi 165873317   ref   ZP 032217921.1 Patt1 gi 262201662   ref   YP 0013272870. gi 27541789   ref   ZP 03971838.1 gi 227505273   ref   ZP 03971838.1 gi 12551827   ref   ZP 03971838.1 gi 19551827   ref   ZP 03971838.1 gi 19551827   ref   ZP 03971838.1 gi 19551827   ref   ZP 0395322.1 gi 19551827   ref   ZP 048593.1 gi 125989429   ref   ZP 048593.1 gi 125989429   ref   ZP 04859584.1 gi 125989429   ref   ZP 0459564.1 gi 123619142   ref   ZP 0459564.1 gi 1238926143   ref   ZP 0405753.1 gi 1238926143   ref   ZP 0405753.1 gi 1261878007   ref   ZP 0604612.1 gi 176044322   ref   YP 002352857. gi 127967351   ref   YP 002352857. gi 12309102   ref   YP 0023568.1 gi 123955411   ref   ZP 0444320.1 gi 1268312   ref   YP 044320.1</pre>		SREAT IN JUTN OF A WY KREATEN JEH SFAM R KREATEN JEH SFAM R KREATEN JEH SFAM R FRAIG HEITMEHEN SREHGE-TISGIRHY YRE-EGE-TISGIRHY YRE-EGE-VOLGVRN YRE-FGE-VOLGVRN YEA-FGE-KILAVRN YEA-FGE-NUSVRH YEA-FGE-NUSVRH YEA-FGE-NUSVRH YEA-FGE-NUSVRH YEA-FGE-NUSVRH YEA-FGE-NUSVRH YEA-FGE-NUSVRH YEA-FGE-KILGRRD YEA-FGE-TIRGR YEA-FGE-KILGRRD YEA-FGE-TIRGR YEA-FGE-KILGVRN YEA-FGE-KILGVRN YEA-FGE-KILGVRN YEA-FGE-KILGVRN YEA-FGE-KILGVRN YEA-FGE-KILGVRN YEA-FGE-KILGVRN YEA-FGE-KILGVRN YEA-FGE-KILGVRN YEA-FGE-KILGVRN YEA-FGE-KILGVRN YEA-FGE-KILGVRN YN KGE-KILGVRN YN KGE-KILAVGRRN	Y Y = I Y E + C = I G Y G = G V E + C G X I G Y G = G V E + C G X V G Y G = G Y
<pre>gi  15613110 ref NP 241413 1] gi  205372356 ref  ZP 03225170.1 gi  163938250 ref  ZP 03225170.1 gi  163938250 ref  ZP 001643134. gi  165873317 ref  ZP 03272870. gi  262201662 ref  YP 003272870. gi  227541789 ref  ZP 03971838.1 gi  227505273 ref  ZP 03971838.1 gi  227505273 ref  ZP 03935322.1 gi  19963453 ref  XP 948593.1  gi  3552802 ref  NP 948593.1  gi  195543]ref  ZP 048593.1  gi  15528429 ref  ZP 048593.1  gi  15528429 ref  ZP 048593.1  gi  1528429 ref  ZP 048593.1  gi  1528429 ref  ZP 0459568.1 gi  23019142 ref  ZP 0459568.1 gi  226558075 ref  ZP 0459568.1 gi  22031046 ref  ZP 0459568.1 gi  22031046 ref  ZP 0457903.1] gi  203031046 ref  ZP 0457903.1] gi  208429 ref  ZP 04575.1] gi  208429 ref  ZP 04575.1] gi  2084059 ref  ZP 04576.1] gi  206900711 ref  ZP 002352857. gi  227957351 ref  ZP 04443200.1] gi  23098102 ref  NP 691568.1] gi  22098102 ref  NP 04443200.1] gi  23098102 ref  ZP 04443200.1] gi  23098102 ref  ZP 04443200.1] gi  23098102 ref  ZP 044701.1] gi  21283703 ref  ZP 044701.1]</pre>	MARSREIGVEN HADOR SINGE GENERGEN UND UDA UND EN DE VENERS T- VE	REBATING     FINGTING     FARATY       KRSAFENELENSFARNY     KRSAFENELENSFARNY       KRSAFENELENSFARNY     FRALQE     FRALQE       TADAFHRELENSFARNY     FRALQE     FRALQENSFARNY       YRR-EGE     FISGIRRHY       YRR-EGE     FISGIRRHY       YRR-EGE     FISGIRRHY       YRR-EGE     FISGIRRHY       YRR-EGE     FISGIRRHY       YEA-FGE     FISGRRHY       YEA-FGE     FISGRRHY       YEA-FGE     FISGRRHY       YEA-FGE     FISGRRHY       YEA-FGE     FISGRRHY       YEA-GGE     FISGRRHY       Y	$ \begin{array}{l} \mathbf{y} = -\mathbf{y} = -\mathbf{y} \in \mathbf{g} \in \mathbf{y} \in \mathbf{y} \in \mathbf{y} \\ \mathbf{y} = -\mathbf{y} = -\mathbf{y} \in \mathbf{g} \in \mathbf{y} \in \mathbf{y} \in \mathbf{y} \\ \mathbf{y} = -\mathbf{z} = -\mathbf{k} \cap \mathbf{g} \in \mathbf{y} \in \mathbf{y} \in \mathbf{y} \in \mathbf{y} \\ \mathbf{y} = -\mathbf{z} = -\mathbf{k} \cap \mathbf{g} \in \mathbf{y} \in \mathbf{y} \in \mathbf{y} \\ \mathbf{y} = -\mathbf{y} \in \mathbf{g} \cap \mathbf{h} \cap \mathbf{g} = \mathbf{y} \in \mathbf{y} \\ \mathbf{y} = -\mathbf{y} \in \mathbf{g} \cap \mathbf{h} \cap \mathbf{h} \\ \mathbf{y} = -\mathbf{y} \in \mathbf{g} \cap \mathbf{h} \cap \mathbf{h} \\ \mathbf{y} = -\mathbf{y} \in \mathbf{g} \cap \mathbf{h} \cap \mathbf{h} \\ \mathbf{y} = -\mathbf{y} \in \mathbf{g} \cap \mathbf{h} \cap \mathbf{h} \\ \mathbf{y} = -\mathbf{y} \in \mathbf{g} \cap \mathbf{h} \cap \mathbf{h} \\ \mathbf{y} = -\mathbf{y} \in \mathbf{g} \cap \mathbf{h} \cap \mathbf{h} \\ \mathbf{y} = -\mathbf{y} \in \mathbf{g} \cap \mathbf{h} \cap \mathbf{h} \\ \mathbf{y} = -\mathbf{y} \in \mathbf{g} \cap \mathbf{h} \cap \mathbf{h} \\ \mathbf{y} = -\mathbf{y} \in \mathbf{g} \cap \mathbf{h} \cap \mathbf{h} \\ \mathbf{y} = -\mathbf{y} \in \mathbf{g} \cap \mathbf{h} \\ \mathbf{y} = -\mathbf{y} = \mathbf{y} = \mathbf{h} \\ \mathbf{y} = -\mathbf{y} = \mathbf{y} = \mathbf{h} \\ \mathbf{y} = -\mathbf{y} = \mathbf{y} = \mathbf{h} \\ \mathbf{y} = -\mathbf{y} = \mathbf{h} \\ \mathbf{y} = -\mathbf{y} = \mathbf{h} \\ \mathbf{y} = -\mathbf{h} \\ \mathbf{h} = \mathbf{h} \\ \mathbf{h} \\ \mathbf{h} \\ \mathbf{y} = -\mathbf{h} \\ \mathbf{h} \\ \mathbf$
gi   15613110   rof   NP 241413   1     gi   205372356   ref   ZP 03225170.1     gi   163938250   ref   YP 001643134.     gi   165873317   ref   ZP 03225170.1     gi   262201662   ref   YP 001643134.     gi   262201662   ref   YP 003272870.     gi   2675173   ref   ZP 03395322.1     gi   25258082   ref   YP 948593.1      gi   37528082   ref   YP 948593.1      gi   37528082   ref   XP 948593.1      gi   170768543   ref   ZP 03305322.1     gi   170768543   ref   XP 948593.1      gi   22556075   ref   XP 03059564.1     gi   225931046   ref   XP 00359754.     gi   228926143   ref   ZP 04657903.1     gi   228926143   ref   ZP 04657903.1     gi   228926143   ref   ZP 0604612.1     gi   2060071   ref   ZP 0022506835.1     gi   206071   ref   YP 002352857.     gi   217967351   ref   YP 002352857.     gi   2255541   ref   XP 0444320.1     gi   2268543   ref   YP 0444320.1     gi   2248703   ref   YP 0444320.1     gi   22385541   ref   ZP 0668.1     gi   2248753   ref   XP 0444320.1     gi   22488750   ref   YP 0444320.1		SREATS IN JUTN OF A MUY READA HREIEVE AM R FRAIDS - EIDDSSPS WSR-HGE-TTSGIRHY YRR-EGE-TTSGIRHY YRR-EGE-VILGYRKN YRR-FGE-VILGYRKN YRR-FGE-NCUSYRRN YEA-FGE-NCUSYRRN YEA-FGE-SIGN YEA-FGE-RIANN YEA-FGE-SIGN YEA-FGE-RIANN YEA-FGE-RIANN YEA-FGE-RIANN YEA-FGE-RIANN YEA-FGE-RIANN YEA-FGE-RIANN YEA-FGE-RIANN YEA-FGE-RIANN YEA-FGE-RIANN YEA-FGE-RIANN YEA-FGE-SAUGYRRN YEA-FGE-YEA YEA-FGE-YEA YEA-FGE-YEA YEA-FGE-YEA YEA-FGE-YEA YEA-FGE-YEA YEA YEA-FGE-RIANN YEA-FGE-YEA YEA YEA-FGE-YEA YEA YEA-FGE-YEA YEA YEA-FGE-YEA YEA YEA YEA YEA YEA YEA YEA	Y =
gi   156:13110   ref NP 241413 1     gi   205372356   ref   ZP 03225170.1     gi   263372356   ref   ZP 03225170.1     gi   163932250   ref   ZP 03225170.1     gi   16587317   ref   ZP 032217921.1     Patt1     gi   262201662   ref   YP 003272870.     gi   262201662   ref   YP 003272870.     gi   262201662   ref   YP 003272870.     gi   26751780   ref   ZP 03971838.1     gi   227505273   ref   ZP 03971838.1     gi   2505273   ref   ZP 03971838.1     gi   15298423   ref   XP 948593.1     gi   177747765   ref   XP 02902996.1     gi   15298429   ref   XP 035075.2     gi   2350282   ref   XP 035075.1     gi   235028429   ref   ZP 04599568.1     gi   25298429   ref   ZP 04599568.1     gi   252926143   ref   ZP 04599568.1     gi   25089071   ref   ZP 06004612.1     gi   251878007   ref   ZP 06025798.1     gi   251878007   ref   ZP 06025795.1     gi   217967351   ref   YP 002352857.     gi   217967351   ref   YP 002352857.     gi   217967351   ref   YP 0123568.1     gi   217967351   ref   YP 0443200.1     gi   217967351   ref   YP 0443200.1     gi   217967351   ref   YP 0443200.1     gi   217967351   ref   YP 04443200.1     gi   21796	MARSREIE GVEVHIAQUENGEVGIVADUENDAUVRÜRDAT UF 	REBAT IN TYTNO PARWYRK       RESATENDIENS FARWYRK       KRSAFENDIENS FARWYRK       KRSAFENDIENS FARWYRK       KRSAFENDIENS FARWYRK       TADAFHREITWEHHWY       FRALOF       SRHGE       VRESAFENDIENS       YRE-EGF       FRALOF       VRE-FGF       VRE-FGF       SRENGENS       YES-IGF       YES-YGC       YES-YGC       YES-IGF       YES-IGF<	$ \begin{array}{c} \mathbf{y} = -\mathbf{y} = -\mathbf{y} \in \mathbf{g} \in \mathbf{y} \in \mathbf{y} \in \mathbf{y} \\ \mathbf{y} = -\mathbf{y} = \mathbf{z} \in \mathbf{g} \in \mathbf{g} \in \mathbf{y} \in \mathbf{y} \\ \mathbf{y} = -\mathbf{z} = \mathbf{z} \in \mathbf{g} \in \mathbf{z} \in \mathbf{y} \in \mathbf{z} \in \mathbf{y} \in \mathbf{z} \\ \mathbf{y} = -\mathbf{z} = \mathbf{z} \in \mathbf{g} \in \mathbf{z} \in \mathbf{z} = -\mathbf{z} \\ \mathbf{y} = -\mathbf{z} \in \mathbf{z} \in \mathbf{z} \in \mathbf{z} \\ \mathbf{z} = -\mathbf{z} = -\mathbf{z} \in \mathbf{z} \in \mathbf{z} \\ \mathbf{z} = -\mathbf{z} = -\mathbf{z} = -\mathbf{z} \\ \mathbf{z} \\ \mathbf{z} = -\mathbf{z} \\ \mathbf{z} = -\mathbf{z} \\ \mathbf{z} \\ \mathbf{z} = -\mathbf{z} \\ \mathbf{z} = -\mathbf{z} \\ \mathbf{z} \\ \mathbf{z} = -\mathbf{z} \\ \mathbf{z} \\ \mathbf{z} = -\mathbf{z} \\ \mathbf{z} \\ \mathbf{z} \\ \mathbf{z} = -\mathbf{z} \\ \mathbf{z} \\ \mathbf{z} = -\mathbf{z} \\ \mathbf{z} \\ \mathbf{z} \\ \mathbf{z} = -\mathbf{z} \\ \mathbf{z} \\ \mathbf{z} \\ \mathbf{z} \\ \mathbf{z} = -\mathbf{z} \\ \mathbf{z} \\ \mathbf{z} \\ \mathbf{z} \\ \mathbf{z} = -\mathbf{z} \\ \mathbf{z} $
gi   15613110   ref   NP 241413   1     gi   205372356   ref   ZP 03225170.1     gi   163938250   ref   YP 001643134.     gi   165873317   ref   ZP 03225170.1     Patt1     gi   262201662   ref   YP 001643134.     gi   262201662   ref   YP 003272870.     gi   262201662   ref   YP 003272870.     gi   262201662   ref   YP 003272870.     gi   227505273   ref   ZP 03971838.1     gi   227505273   ref   ZP 03971838.1     gi   25258082   ref   NP 59829.1     gi   37528082   ref   NP 59829.3     gi   37528082   ref   NP 59829.3     gi   37528082   ref   NP 5948593.1     gi   37528082   ref   NP 50835.2     gi   220931046   ref   NP 50835.2     gi   2289543   ref   ZP 045578.1     gi   226900711   ref   ZP 002552857.3     gi   206900711   ref   XP 002552857.3     gi   22955411   ref   ZP 0443220.1     gi   22955411   ref   ZP 0443220.1     gi   22955411   ref   ZP 0443220.1     gi   22955431   ref   XP 044320.1     gi   2296313   ref   NP 646791.1     gi   2296313   ref   NP 044320.1     gi   22637383   ref   NP 044320.1		SREATS IN TYTM OPAR W REATER IN THIS PAR SAME TADAFHREIFYNEHAR Y TADAFHREIFYNEHAR Y FRAIGT - TSGIRHY YRR-EGF-TSGIRHY YRR-EGF-TSGIRHY YRR-FGF-YOLGYRKY YRR-FGF-YOLGYRKY YRR-FGF-NUSYRRY YEA-FGF-NUSYRRY YEA-FGF-NUSYRRY YEA-FGF-NUSYRRY YEA-FGF-NUSYRRY YEA-FGF-SGIRHY YEA-GGIRHY YA-GGIRHY YA-	Y
<pre>gi   15613110  rof NP 241413 1 gi   205372356  ref   ZP 03225170.1 gi   163932250  ref   ZP 03225170.1 gi   163932250  ref   ZP 03225170.1 gi   165873317  ref   ZP 03217921.1 Patt1 gi   262201662  ref   YP 001272870. gi   21741789  ref   ZP 03971838.1 gi   227505273  ref   ZP 03971838.1 gi   225505273  ref   ZP 03971838.1 gi   25505273  ref   ZP 03971838.1 gi   25505273  ref   ZP 03971838.1 gi   37528082  ref   NP 948593.1 gi   37528082  ref   NP 948593.1 gi   37528082  ref   ZP 0459768.1 gi   37528082  ref   ZP 0459568.1 gi   152989429  ref   ZP 0459568.1 gi   255658075  ref   ZP 0459568.1 gi   25658075  ref   ZP 0459578.1 gi   256870071   ref   ZP 0604612.1 gi   71644322  ref   YP 0755.1 gi   25697751   ref   ZP 0604612.1 gi   27697351   ref   ZP 0457903.1 gi   2555541   ref   ZP 0457903.1 gi   226971   ref   ZP 045758.1 gi   22555541   ref   ZP 0465781.1 gi   2069711   ref   ZP 045758.1 gi   22955541   ref   ZP 044320.1 gi   22468312   ref   ZP 044320.1 gi   226871831   ref   ZP 0443220.1 gi   2468769   ref   ZP 04578747.1 gi   246859   ref   ZP 04578747.1 gi   246859   ref   ZP 04578747.1 gi   2464255   ref   ZP 045227340.1 gi   2464255   ref   ZP 0452727340.1 gi   2464255   ref   ZP 0452727340.2 gi   227468569   ref   ZP 04578747.1 gi   2464255   ref   ZP 04527340.2 gi   227340.3   ref   ZP 04578747.1 gi   261403085   ref   ZP 04527340.3   ref   ZP 04578747.1 gi   261403085   ref   ZP 02327340.3   ref   ZP 04578747.1 ] gi   261403085   ref   ZP 045787477.1 ] gi   261403085   ref   ZP 02327340.1 ]</pre>		REBAT     INFUTNO     PARMY       RRAFE     INFUTNO     PARMY       REAF     INFUTNO     PAR	Y G = - Y E G Q Y Q Y     Y       Y Y - E K Q Q Y Q Y     Y       Y E K D G Y Y Q Y     Y       Y P S G A D A P T Y     Y       Y P S G A D A P T Y     Y       Y P S G A D A P T Y     Y       Y - P S G A D A P T Y     Y       Y P S G A D A P T Y     Y       Y - P S G A D A P T Y     Y       Y - P S G A D A P T Y     Y       Y - P S G A D A P T Y     Y       Y - P S G A D A P T Y     Y       Y - P S G A D A P T Y     Y       Y - P S G A D A P T Y     Y       Y - P S G A D A P T Y     Y       Y - P S G A D A P T Y     Y       Y P A T G K E D A P T Y     Y       Y P A T G K E D A P T Y     Y       Y - D N C D A P T Y     Y       Y - D N C D A P T Y     Y       Y - D N C D A P T Y     Y       Y D N C D A P T Y     Y       Y D N C D A P T Y     Y       Y D N N E N A P - Y Y     Y       Y D N N E N A P - Y Y     Y
<pre>ilistiatio ref NP 241413 11 gi 205372356 ref YP_00225170.1 gi 163938250 ref YP_001643134. gi 165873317 ref YP_001643134. gi 165873317 ref YP_00217921.1 Patt1 gi 226201662 ref YP_003272870. gi 2474789 ref YP_003971838.1 gi 227505273]ref YP_0395322.1 gi 12578082 ref YP_948593.1  gi 13528082 ref YP_948593.1  gi 13528082 ref YP_948593.1  gi 1707655ref YP_00350715. gi 23758082 ref YP_00350715. gi 2389429 ref YP_00350715. gi 2389429 ref YP_00359784.1 gi 228956143 ref ZP_04599568.1 gi 228956143 ref ZP_0459578.1  gi 2389525441]ref ZP_045578.1  gi 23091042 ref YP_07575.1] gi 2309102 ref YP_075578.1] gi 2309102 ref YP_014701.1] gi 229370316[YP_0145780.1] gi 220931046]ref ZP_04657903.1] gi 23098102 ref YP_014701.0]. gi 23098102 ref YP_014701.1] gi 223955411 ref ZP_04578.1] gi 22493731 ref ZP_047871.1] gi 2248569 ref PP_042524526.1] gi 22485783 ref ZP_04578.1] gi 22485783 ref ZP_04578.1] gi 2248561 ref ZP_04578.1] gi 2248561 ref ZP_04578.1] gi 2248561 ref ZP_04578.1] gi 224857831 ref ZP_0457878.1] gi 224537831 ref ZP_0457878.1] gi 2245377831 ref ZP_0457878.1] gi 224537783</pre>	MARSREIG VENH TADO ENGLIG FUNDEUDAU DU VENDEUDAU STUDIE 	S R E R I M UT M O PAR W R S P E M F T M O PAR W T A D A F H R I E V E H A H V T A D A F H R I E V E H A H V T A D A F H R I E V E H A H V T A D A F H R I V M E H A H V M S A H G I T S G T R H V M S A H G I V I V M V M K A H M V W S A F G I V I V M K A H M V W S A F G I N V M V M K A H M V W S A F G I N V M V M K A H M V M K A H M V W S A F G I N V M K A H M V M K A H M V M K A H M V M K A H M V M K A H M V M K A H M V M K A H M V M K A H M V M K A H M V M K A H M K M K M K M K M K M K M K M K M K M	Y Y Y Y Y Y Y Y Y Y Y Y
<pre>i i 5613110   rof NP 241413 1 gi 205372356   ref   ZP 03225170.1 gi 163372356   ref   ZP 03225170.1 gi 163372356   ref   ZP 03225170.1 gi 165873317   ref   ZP 03217921.1 Patt1 gi 262201662   ref   YP 001643134. gi 262201662   ref   YP 001372870. gi 27505273   ref   ZP 03375322.1 gi 12551827   ref   ZP 03375322.1 gi 13551827   ref   ZP 03375322.1 gi 135528082   ref   XP 948593.1  gi 135528082   ref   XP 948593.1  gi 137528082   ref   XP 03929.1] gi 137528082   ref   XP 039537838.1] gi 2751827   ref   ZP 048593.1] gi 135528082   ref   XP 048593.1] gi 125658075   ref   ZP 048593.1] gi 125658075   ref   ZP 04959568.1 gi 220931046   ref   XP 002507954. gi 225658075   ref   ZP 04657903.1] gi 1261878007   ref   ZP 04657903.1] gi 1261878007   ref   ZP 04657903.1] gi 1261878007   ref   ZP 04587.3] gi 127967351   ref   XP 035887.1] gi 127967351   ref   XP 03252887. gi 12395131   ref   XP 044320.1 gi 1268591071   ref   ZP 0443200.1 gi 1268591071   ref   ZP 0443200.1 gi 12637783   ref   XP 04678747.1] gi 12637783   ref   XP 047701.1] gi 12637783   ref   XP 042245326. gi 120197058   ref   XP 0423245326. gi 120197058   ref   XP 0437847.1] gi 12637783   ref   XP 0437847.1] gi 12637783   ref   XP 0437847.1] gi 12637783   ref   XP 0437847.1] gi 12637783   ref   XP 0437847.1] gi 122637783   ref   XP 0437847.1] gi 122637783   ref   XP 0437847.1] gi 122637783   ref   XP 0437847.1] gi 122637783   ref   XP 0437847.1] gi 12263783   ref   XP 0437847.1] gi 12263783   ref   XP 0437847.1] gi 12263783   ref   XP 0437847.1] gi 122637783   ref   XP 0437847.1] gi 12263783   ref   XP 0437847.1] gi 12263784   ref   XP 0437847.1] gi 122637847.1] gi 1263783   ref   XP 0457847.1] gi 12637847.1] gi 12647847.1] gi 12647847.1] gi 12647847.1] gi 12796758   ref   YP 047787.1] gi 12796758   ref   YP 0477877.1] gi 126178707578   ref   YP 047777.1] gi 126178707787.1] gi 126177777777</pre>	MARSREIGGVENT AUX MARKERIGT VUD UD AVRUGEON TVD	SRENT IN TYTN OPAN W RESATEN STANDATH STANDATH RESATEN STANDATH STANDATH RESATEN STANDATH STANDATH FRAIDT HELT NEH BHUY FRAIDT HELT NEIGHT STANDATH FRAIDT HELT NOUL STANDATH FRAIDT HELT NULL STAND	Y
<pre>i   15613110   rof NP 241413   1 gi   205372356   ref   ZP 03225170.1 gi   163938250   ref   ZP 03225170.1 gi   163873371   ref   ZP 022717921.1 Patt1 gi   262201662   ref   YP 003272870. gi   21637471 ref   YP 003272870. gi   2127541789   ref   YP 003371838.1 gi   227503273   ref   ZP 0335322.1 gi   13951827   ref   NP 0395322.1 gi   13951827   ref   NP 048593.1 gi   13951827   ref   NP 048593.1 gi   13758082   ref   NP 048593.1 gi   13774765   ref   NP 048593.1 gi   220931046   ref   NP 0459586.1 gi   220931046   ref   NP 00350715. gi   2289429   ref   YP 00357783.1 gi   22697301   ref   YP 076750.1 gi   25896075   ref   YP 076750.1 gi   27044322   ref   YP 076750.1 gi   23098102   ref   YP 076756.1 gi   23098102   ref   YP 014701.1 gi   22493783   ref   YP 0444320.1 gi   226937783   ref   YP 04771.1 gi   226937783   ref   YP 035278.1 gi   22093778   ref   YP 04771.1 gi   22637783   ref   YP 032528578.1 gi   22093778   ref   YP 03252878.1 gi   22093778   ref   YP 03252878.1 gi   22093778   ref   YP 014701.1 gi   2264090851   ref   YP 03252878.1 gi   22093778   ref   YP 0147071.1 gi   2248569   ref   YP 0147071.1 gi   2248569   ref   YP 0147071.1 gi   2248569   ref   YP 0147071.1 gi   2257783   ref   YP 0147071.1 gi   221795058   ref   YP 00327380.7 gi   163762778   ref   YP 00327789.5 gi   1637632   ref   YP 00327789.5 gi   1637632   ref   YP 00327789.5 gi   16376331   ref   YP 00309789.5 gi   16376331   ref   YP 0014701.5 gi   16376331   ref   YP 00309789.5 gi   16376331   ref   YP 00309789.5 gi   16376331   ref   YP 00309789.5 gi   16376331   ref   YP 00309789.5 gi   16376331   ref   YP 00309789.5 gi</pre>	MARSREIGVEN HADOR MARKER ING WOLD DA WINDERT - UP	S REF IN TYTN O PAR W REF IN TYTN O PAR W REF IN THE IF IN FRAME TADATHR IF W FRAME TADATHR IF W FRAME TADATHR IF W FRAME FRAME REF IN TO REF INT REF	Y

Fig. 1 Multiple sequence alignment of Patt1 amino acid sequence (residues 50–237) and representative sequences of significant similarity to Patt1. GNAT family conserved sequence motifs [11] are discernible

in the region of predicted disorder and is likely to represent "intrinsically disordered" locally stable alpha helix in the lack of defined tertiary structure.

#### Fold recognition and Patt1 structure modeling

Fold recognition methods implemented in GeneSilico Metaserver were used to search within PDB database for protein structures solved by X-ray crystallography and NMR. The structure scoring by fold recognition servers provides means of selection of suitable template structures. Fold recognition servers returned many structures that can be used as possible templates in homology modeling. The templates were in most cases members of a newly discovered GNAT subfamily of pita-like proteins. The choice was based on sequence similarity and structure scoring by fold recognition servers. The structure of RimI [30] was chosen as a modeling template. The Patt1 model was constructed with use of the MODELLER program.

#### Patt1 model assessment

The Patt1 model was scored by ProQ and MetaMQAPII methods. The ProQ LGscore: 2.082 and MaxSub: 0.275 indicated that the model was of sufficient quality "fairly good model", and can be used in subsequent stages of analysis. MetaMQAPII confirmed correctness of the model.

# Patt1 molecular dynamics

The binding site of the peptide and AcCoA is located at the catalytic center of the enzyme, surrounded by B-sheet elements  $\beta 6$  and  $\beta 7$ . The extended loop between these structural elements showed highest conformational flexibility during



Fig. 2 a Ensemble of structures from MD simulation representing stages of the model after 4, 6, 8, 10 and 12 ns. Colors corresponds to RMSD values. b Plot of RMSD value averaged for C-alpha atoms over time

simulation. This behavior corresponds with available literature data for RimI where  $\beta 6$ - $\beta 7$  loop movements correlate with acetyltransfer reaction. The protein structure was stable during simulation, no unfolding of tertiary structure was observed (Fig. 2).

# Patt1 structure analysis

The region that comprises GNAT fold domain is located between 59 and 220 amino acid residue. Patt1 reveals classical

GNAT fold with the central region of beta sheet surrounded by alpha helices. The sequence of Patt1 structural elements is:  $\beta_{1-\alpha_{1-\alpha_{2}-\beta_{2-\beta_{3-\beta_{4-\alpha_{3-\beta_{5-\alpha_{4-\beta_{6-\beta_{7}}}}}}}$  (Fig. 3). In the core region, the enzyme active site is located at the edge of the central beta sheet, surrounded by  $\alpha_{3}$  and  $\alpha_{4}$  helices in the region of Acetyl-CoA binding. The  $\beta_{6-\beta_{7}}$  sheet structural elements along with the loop between  $\alpha_{1}$  and  $\alpha_{2}$  helices enclose the active site in the region of acetylated peptide binding. The characteristic feature of GNAT family proteins, the V-like splay shape between the  $\beta_{4-\beta_{5}}$  (the specific region



<sup>(1)</sup> <sup>(2)</sup> <sup>(3)</sup> <sup>(3)</sup> <sup>(2)</sup> <sup>(3)</sup> <sup>(4)</sup> <sup>(4)</sup>

Fig. 3 Overview of the Patt1 structure. Schematic ribbon representation featuring sheets and helices. The second picture is rotated by  $90^{\circ}$  along x axis into image plane, giving impression of "top

view" over the edge of the central  $\beta$ -sheet. The  $\beta$ -like structural element located between  $\beta 6$ - $\beta 7$  sheets is unstable and dissolves during MD simulation into extended loop

of active site where acetvltransfer reaction takes place) structural elements is also present in the Patt1 model. The high quality of the model allows detailed predictions about the specific amino acid residues function to be made. The sequence region "RRKGLG" spanning between 147 and 152 amino acid residues in Patt1 sequence was shown to be the GNAT signature motif Arg/Gln-X-X-Gly-X-Gly/Ala responsible for recognition and binding of CoA [13]. In the model it is located in the loop between  $\beta 4-\alpha 3$  and at the beginning of  $\alpha$ 3 helix. Particularly interesting are other interactions with CoA in Patt1 structure. In the  $\alpha$ 4 helix there are two highly conserved phenylalanine residues present: Phe185 and Phe186 that are directed to CoA substrate. With the next highly conserved Phe192 residue located within  $\alpha$ 4- $\beta$ 6 loop that follows  $\alpha 4$  helix they form a system of stacking rings. CoA adenine rings are trapped in contacts with the plane of the Phe185 ring. These interactions contribute to proper positioning of the ligand. Structural alignment of Patt1 and RimI reveals that spatial localization of Phe186 residue in Patt1 corresponds to Tyr115 in RimI structure. Tyr115 (RimI) has been assigned a role of the active site acid [30]. The interactions with stacking rings systems have been shown to act as an alternative to general acid catalysis [31]. Therefore, it is proposed that this spatial arrangement of conserved phenolic rings should take part in acetyltransfer reaction (Fig. 4). In the structure of RimI the amide backbone of the residue Ile69 has a role of stabilizer of polarization of acetyl group in the process in which the tetrahedral intermediate is formed during acetyltransfer reaction. This position is occupied by a conserved Val140 residue in Patt1. In the structure of Patt1 in the



Fig. 4 Spatial localization of key residues in Patt1 active site. The CoA ligand was copied from RimI structure

active site there is conserved Cysteine 137 residue located with sulfhydryl group directed to sulfhydryl group of CoA (Fig. 4). The presence of cysteine in the active site might suggest its involvement in reaction mechanism. Structural alignment of Patt1 model with yeast Esa1 structure [32] was performed. Esal is a histone acetyltransferase that belongs to the MYST subfamily. In Esa1 the strictly conserved Cys304 residue is responsible for the common acetvlation reaction mechanism (that proceeds through an acetyl-cysteine enzyme intermediate) [33]. Both Patt1 and Esa1 perform acetylation of histone H4. The structural alignment reveals that both Patt1 Cys137 and Esa1 Cys304 residues are located in the active site, at the beginning of  $\beta$ 4 sheet. The Glu139 residue is localized in close proximity in the Patt1 model (within the  $\beta$ 4 sheet structural element) which has been assigned a role as an active site base [13]. Conserved residues are located in the region of specific interactions with the peptide that undergoes acetylation. These residues surround the peptide substrate and position it prior to acetyltransfer. Among them there are residues located in the  $\alpha 1$ - $\alpha 2$  loop. This loop is extended in comparison with RimI structure, due to probable insertion event. The most highly conserved residue in this region-Trp92 is engaged in stacking interactions with conserved Tyr85. This specific positioning increases the strength of hydrogen bonding interactions made by Tyr85 at the recognition interface of Patt1. The conserved Glu86 residue is directed into the space of peptide binding. In the loop that extends between  $\beta 6$ - $\beta 7$  sheet elements a sequence 'GCCG' (Gly203, Cys204, Cys205, Gly206) is located. Glycine residues facilitate conformational flexibility in this region. The presence of hydrophobic amino acid residue in the loop that is exposed to solvent indicates that the loop constitutes a part of recognition interface. In the structure of RimI in this region there are strictly conserved Tyr129 and Tyr130 residues that make contacts with the ligand peptide backbone [30]. Movements of these residues in RimI structure accompanies critical steps of acetyltransfer reaction. Met81(Patt1) is at the place of His23(RimI). Its proposed function, as in the case of RimI analog is to increase affinity to CoA by hydrophobic interactions.

# Conclusions

The all-atom tertiary structure of a novel human histone acetyltransferase Patt1 was modeled by means of theoretical methods. The model was assessed with the state-of-the-art methods for protein structure validation, and subjected to 12 ns molecular dynamics simulation. The integrity of structure was retained. The model allowed to infer sequence-structure-function relationships of Patt1. The key residues identified as involved in acetyltransefer reaction were: Phe185, Phe186, Phe192, and Cys137. Patt1 promotes

apoptosis in human hepatocellular carcinoma cell lines. The availability of effective therapy is strongly dependent on structural data of proteins. Three-dimensional molecular structure can be used as a map in performing protein design experiments and evaluation of their results. It also opens the possibility of rational drug design and can help in the development of personalized medicine.

Acknowledgments All calculations that required the use of high performance computing resources were carried out at the Academic Computer Center in Gdańsk.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## References

- 1. Loidl P (2001) Introduction: assembly, remodeling and modification of chromatin. CMLS Cell Mol Life Sci 58:663–664
- Wu J, Grunstein M (2000) 25 years after the nucleosome model: chromatin modifications. Trends Biochem Sci 25(12):619–623
- Strahl BD, Allis CD (2000) The language of covalent histone modifications. Nature 403(6765):41–45
- Polevoda B, Sherman F (2002) The diversity of acetylated proteins. Genome Biol 3(5):reviews0006
- Marmorstein R (2001) Structure of histone acetyltransferases. J Mol Biol 311:433–444
- Marmorstein R (2001) Structure and function of histone acetyltransferases. CMLS Cell Mol Life Sci 58:693–703
- Timmermann S, Lehrmann H, Polesskaya A, Harel-Bellan A (2001) Histone acetylation and disease. CMLS Cell Mol Life Sci 58:728–736
- Di Gennaro E, Bruzzese F, Caraglia M, Abruzzese A, Budillon A (2004) Acetylation of proteins as novel target for antitumor therapy: review article. Amino Acids 26:435–441
- Dyda F, Klein DC, Hickman AB (2000) GCN5-related Nacetyltransferases: a structural overview. Annu Rev Biophys Biomol Struct 29:81–103
- Vetting MW, de Carvalho LPS, Yu M, Hedge SS, Magnet S, Roderick SL, Blanchard JS (2005) Structure and functions of the GNAT superfamily of acetyltransferases. Arch Biochem Biophys 433(1):212–226
- Neuwald AF, Landsman D (1997) GCN5-related histone Nacetyltransferases belong to a diverse superfamily that includes the yeast SPT10 protein. Trends Biochem Sci 22(5):154–155
- Brownell JE, Zhou J, Ranalli T, Kobayashi R, Edmondson DG, Roth SY, Allis CD (1996) Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. Cell 84(6):843–851
- 13. Liu Z, Liu Y, Wang H, Ge X, Jin Q, Ding G, Hu Y, Zhou B, Chen Z, Ge X, Zhang B, Man X, Zhai Q (2009) Patt1, a novel protein acetyltransferase that is highly expressed in liver and downregulated in hepatocellular carcinoma, enchances apoptosis of hepatoma cells. Int J Biochem Cell Biol 41(12):2528–2537
- 14. Wheeler DL, Barret T, Benson DA, Bryant SH, Canese K, Chetvernin V, Church DM, Dicuccio M, Edgar R, Federhen S, Feolo M, Geer LY, Helmberg W, Kapustin Y, Khovayko O, Landsman D, Lipman DJ, Madden TL, Maglott DR, Miller V,

Ostell J, Pruitt KD, Schuler GD, Schumway M, Sequeira E, Sherry ST, Sirotkin K, Souvorov A, Starchenko G, Tatusov RL, Tatusova TA, Wagner L, Yashenko E (2008) Database resources of the National Center for Biotechnology information. Nucleic Acids Res 36(Database issue):D13–D21

- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25(17): 3389–3402
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32(5):1792–1797, Print 2004
- Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinforma 5:113
- Marchler-Bauer A, Bryant SH (2004) CD-search: protein domain annotations on the fly. Nucleic Acids Res 32(Web Server issue): W327–W331
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98
- Galtier N, Gouy M, Gautier C (1996) SEAVIEW and PHYLO\_WIN: two graphic tools for sequence alignment and molecular phylogeny. Comput Appl Biosci 12(6):543–548
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25(24):4876–4882
- Page RD (1996) TreeView: an application to display phylogenetic trees on personal computers. Comput Appl Biosci 12(4): 357–358
- Kurowski MA, Bujnicki JM (2003) GeneSilico protein structure prediction meta-server. Nucleic Acids Res 31(13):3804– 3807
- Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 234(3):779–815
- Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-Pdb viewer: an environment for comparative protein modeling. Electrophoresis 18(15):2714–2723
- Pawłowski M, Gajda JM, Matlak R, Bujnicki JM (2008) MetaMQAP: a meta-server for the quality assessment of protein models. BMC Bioinforma 9:403
- Wallner B, Elofsson A (2006) Identification of correct regions in protein models using structural, alignment, and consensus information. Protein Sci 15(4):900–913
- Wallner B, Elofsson A (2003) Can correct protein models be identified? Protein Sci 12(5):1073–1086
- Case DA, Cheatham TE 3rd, Darden T, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling C, Wang B, Woods RJ (2005) The Amber biomolecular simulation programs. J Comput Chem 26(16):1668– 1688
- Vetting MW, Bareich DC, Yu M, Blanchard JS (2008) Crystal structure of RimI from Salmonella typhimurium LT2, the GNAT responsible for N(alpha)-acetylation of ribosomal protein S18. Protein Sci 17(10):1781–1790
- Versées W, Loverix S, Vandemeulebroucke A, Geerlings P, Steyaert J (2004) Leaving group activation by aromatic stacking: an alternative to general acid catalysis. J Mol Biol 338(1):1–6
- 32. Yan Y, Barlev NA, Haley RH, Berger SL, Marmorstein R (2000) Crystal structure of yeast Esal suggests a unified mechanism for catalysis and substrate binding by histone acetyltransferases. Mol Cell 6(5):1195–1205
- 33. Yan Y, Harper S, Speicher DW, Marmorstein R (2002) The catalytic mechanism of the ESA1 histone acetyltransferase involves a selfacetylated intermediate. Nat Struct Biol 9(11):862–869