Heliyon 8 (2022) e12203

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

In vitro and *in silico* study of an exclusive insertion in the nicotinamide/ nicotinate mononucleotide adenylyltransferase from *Leishmania braziliensis*



Lesly Johanna Ortiz-Joya^a, Luis Ernesto Contreras Rodríguez^a, Rodrigo Ochoa^b, María Helena Ramírez Hernández^{a,*}

^a Laboratory of Basic Research in Biochemistry. Faculty of Sciences, National University of Colombia, 111321, Bogota, Colombia
^b Biophysics of Tropical Diseases, Max Planck Tandem Group, University of Antioquia, 050010, Medellin, Colombia

ARTICLE INFO

Keywords: Leishmania NAD NMNAT Deletional mutant Molecular dynamics

ABSTRACT

The intracellular parasite *Leishmania braziliensis* is the causal agent of cutaneous and mucocutaneous leishmaniasis, a group of endemic diseases in tropical regions, including Latin America. New therapeutic targets are required to inhibit the pathogen without affecting the host. The enzyme nicotinamide/nicotinate mononucleotide adenylyltransferase (NMNAT; EC: 2.7.7.1/18) is a potential target, since it catalyzes the final step in the biosynthesis of nicotinamide adenine dinucleotide (NAD⁺), which is an essential metabolite in multiple cellular processes. In this work, we produced and evaluated the catalytic activity of the recombinant protein $6His\Delta_{241}$. $_{249}LbNMNAT$ to study the functional relevance of the exclusive insertion present in the enzyme of *L. braziliensis* (LbNMNAT), but absent in the primary structure of human NMNATs. Our results indicate that the 241-249insertion constitutes a structural element that connects the protein structure Rossmann topology with the carboxyl-terminal domain of the enzyme. The removal of this region drastically decreases the solubility, and enzymatic activity of the recombinant, causing its inactivation. Molecular dynamics simulations were carried out with the wild-type and truncated enzymes to verify additional changes in their stability, which indicated a better stability in the wild-type protein. These findings constitute an initial step to identify a new inhibition mechanism for the development of focused pharmacological strategies on exclusive insertions from the LbNMNAT protein.

1. Introduction

Most Latin American countries are affected by infectious tropical diseases, with many of them associated with parasites and classified by the World Health Organization as neglected diseases. Among these diseases we have leishmaniasis, a pathology caused by *Leishmania* parasites, which is considered a public health problem due to its impact and intrinsic molecular, clinical, and epidemiological complexity [1]. Annually, an incidence of approximately 1.3 million new clinical cases is reported with a prevalence in 98 countries, including Colombia, and occurring in the three clinical forms: cutaneous, mucosal and visceral leishmaniasis [2]. For the control and treatment of leishmaniasis, the available drugs are expensive, toxic, and often require long periods of supervised therapy [3]. In addition to this, several publications highlight parasite resistance [4, 5, 6], and there are currently no effective vaccines [7].

Given the lack of specific and less aggressive drugs, it is necessary to discover new molecular targets and new therapeutic agents to aid in the prevention and treatment of leishmaniasis [8]. Various investigations indicate that the enzymes involved in the metabolism of NAD⁺ are attractive targets for the development of drugs against a variety of human diseases [9]. Therefore, our work focuses on the NMNAT, an enzyme that catalyzes the condensation of nicotinamide mononucleotide (NMN) or nicotinic acid mononucleotide (NaMN) with the adenylyl group of ATP to form NAD⁺ or NaAD (nicotinic acid adenine dinucleotide) [10].

In the case of *L. braziliensis*, we have carried out the biochemical and functional characterization of the only reported NMNAT (LbNMNAT) [11, 12]. However, and to expand the characterization, this work focuses on the differences at the structural level between the parasite protein and human isoenzymes. Therefore, we evaluated *in vitro* the participation of the specific insertion 241–249 found exclusively on the LbNMNAT protein through the design and study of its respective deletional mutant. We also modeled the enzyme structure with and without the insertion, and simulations were run to study the protein intrinsic flexibility [13, 14]. Currently, the functional relevance of this insertion on LbNMNAT and its involvement in the enzymatic activity is unknown. For this reason, the

* Corresponding author. E-mail address: mhramirezh@unal.edu.co (M.H. Ramírez Hernández).

https://doi.org/10.1016/j.heliyon.2022.e12203

Received 31 May 2022; Received in revised form 29 July 2022; Accepted 30 November 2022

2405-8440/© 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).





functional and structural analysis of the parasite's enzyme and human NMNATs becomes a necessary step towards the search for potential selective inhibitors against LbNMNAT and/or modulators of NAD⁺ metabolism.

2. Materials and methods

2.1. Sequence analysis of the LbNMNAT protein

Bioinformatic analysis of LbNMNAT was performed using the sequence available in UniProtKB with access number A4H990. Physicochemical parameters were calculated with the ProtParam tool available on the ExPASy server [15]. The NetPhos 3.0 tool [16] was used to search for possible phosphorylation sites. To compare the sequence with those NMNATs from human and species of the genus *Leishmania*, a multiple alignment was performed with the ClustalW and Muscle algorithms incorporated in the CLC Genomics Workbench 22.0.1 (https://digitalinsigh ts.qiagen.com/) [17]. The sequences of the three human isoenzymes 1–3 (NCBI access codes: NP_001284707.1, NP_055854.1 and NP_001307440.1, respectively), were contrasted with the sequences of the NMNATs of the *L. braziliensis, L. infantum, L. major, L. donovani* and *L. mexicana* species (NCBI access codes: XP_001563913.1, XP_001464802.1, XP_001682392.1, XP_003860000.1, XP_003874000.1, respectively).

2.2. Expression and purification of the recombinant proteins 6His-LbNMNAT and 6His- $\Delta_{241-249}$ LbNMNAT

The expression vector pQE30- $\Delta_{241-249}$ LbNMNAT was obtained by directed mutagenesis using PCR with the Phusion site direct mutagenesis kit (Finnzymes), following the manufacturer's instructions. As template DNA we used the plasmid pQE30-LbNMNAT [18]. The designed primers (forward: 5' TCC GTG CCC GAC GAC ACA TCC TCG 3', and reverse: 5' GGG GGC AGC AGT AGA GGA TAA CAT 3') flanked the deletion site, guaranteeing that nucleotides 721-747 of the LbNMNAT sequence are not included in the amplicon. The pQE30- $\Delta_{241-249}$ LbNMNAT vector was verified by sequencing and used to transform chemically competent Escherichia coli SHuffle T7 Express cells (New England Biolabs). 200 mL of LB culture medium supplemented with ampicillin (100 μ g/mL) at an OD_{600nm} of 0.65 was induced with 1.0 mM of IPTG (isopropyl β -D-1-thiogalactopyranoside) at 24 °C for 16 h 180 rpm. The culture was centrifuged at 4,000 g for 10 min at 4 °C. The cell pellet obtained was resuspended in lysis buffer (50 mM NaH₂PO₄ pH 7.5, 500 mM NaCl, 15 mM Imidazole) in the presence of a 1:300 protease inhibitor cocktail (Sigma P8340), and lysozyme at a final concentration of 1.0 mg/mL, and incubated at 4 °C for 45 min. Subsequently, the sample was lysed by sonication with pulses of 15 s, pauses of 15 s, and amplitude of 50%, for 5 min using the Ultrasonic Processor VCX130 equipment, with a 6 mm diameter probe. Finally, the bacterial lysate was centrifuged at 10,000 g for 20 min at 4 °C.

Purification was performed by nickel affinity chromatography. The soluble extract obtained from cell lysis was incubated for 1 h on ice under constant agitation with Ni-NTA resin (Quiagen), previously equilibrated with lysis buffer. The protein was eluted from the column using elution buffer (50 mM NaH₂PO₄ pH 7.5, 500 mM NaCl, 300 mM Imidazole). The eluate was dialyzed against one liter of buffer (25 mM NaH₂PO₄ pH 7.5, 500 mM NaCl, 0.5 mM DTT) O.N at 4 °C, using a membrane with a nominal molecular weight limit of 10 kDa. The partial purification of the recombinants was observed by means of 12% SDS-PAGE and the quantification was carried out with the Bradford method using BSA as standard [19].

2.3. Western blotting

For western blotting, 6His-LbNMNAT and 6His- $\Delta_{241-249}$ LbNMNAT proteins were separated on a 12% SDS-PAGE, blotted onto PVDF membrane at 200 mA for 2 h in transfer buffer (10 mM Tris-HCl pH 8.0, 0.2 M

glycine, 10 % (v/v) methanol). The membrane was blocked at room temperature for 2 h with 5 % non-fat milk in TBS-T buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% (v/v) Tween 20). Three washes were performed with TBS-T each for 10 min. The detection of the 6His tag fusion proteins was performed by incubating the membrane with the primary antibody anti-6His (1:5000) (Abcam) and anti-mouse IgG secondary antibody conjugated to biotin (1:10000) (Sigma). Chromogenic detection of complexes was performed with alkaline phosphatase-conjugated streptavidin (1:3000) (Promega) and the substrates nitroblue tetrazolium (NBT) (50 mg/ml), and 5-bromo-4-chloro-3'-indolyl phosphate (BCIP) (50 mg/ml) (Promega) in substrate buffer (100 mM Tris-HCl pH 9.0, 150 mM NaCl, 1 mM MgCl₂).

2.4. Evaluation of the enzymatic activity of recombinant proteins by coupled assays

The enzymatic activity of the recombinant protein was determined through a coupled spectrophotometric assay, based on the reduction of NAD⁺ by the action of the enzyme alcohol dehydrogenase (ADH). In this assay, the activity of the enzyme is measured by the increase in absorbance at 340 nm, due to the reduction of NAD⁺ to NADH [20]. To perform the reactions, we made an initial mixture containing 40 mM ethanol, 25 mM of HEPES-KOH buffer at pH 7.5, 20 mM MgCl₂, 2 mM NMN (Sigma), 0.2 U ADH (Sigma) and between 2.5-10 μ g of the samples to evaluate. The reaction was started by adding 1.0 mM ATP (Sigma) recording the absorbance values every 10 s at 340 nm, for 10 min. We performed each reaction in triplicate in the Jenway 7315 spectrophotometer. We used the partially purified wild-type protein (6His-LbNM-NAT) as the activity control.

2.5. Evaluation of the enzymatic activity of NMNAT by direct assays

In the direct assays, the synthesis of NAD⁺ by the enzyme to be evaluated was analyzed by reverse-phase high-performance liquid chromatography (RP-HPLC). In this case, each protein (6His-LbNMNAT and 6His- $\Delta_{241\cdot249}$ LbNMNAT) was incubated at 37 °C for 30 min with a reaction mixture containing 40 mM HEPES pH 7.8, 10 mM MgCl₂ and 1.25 mM ATP. The reaction was quenched with 1.2 M HClO₄ and neutralized with 1M K₂CO₃. The dinucleotide obtained as a product of the reaction was analyzed in the Agilent 1200 Series chromatograph on a C18 reverse phase column (100 cm long x 4.6 mm internal diameter and 5 µm particle size, Phenomenex), at room temperature with a flow rate of 1.0 ml/min, using buffer A (10 mM NaH₂PO₄ pH 6.0) and buffer B (methanol) in the following elution gradient: 2 min buffer A (100%), 4 min buffer A:B (80 %–20 %), 11 min of buffer A:B (75 %: 25 %) and 10 min of buffer B (100 %). The absorbance of the analytes was monitored at 260 nm.

2.6. Modeling and simulation

The three-dimensional models of LbNMNAT with and without the 241–249 insertion were generated by threading using the I-TASSER server [21, 22, 23]. The models were assessed in terms of the dihedrals distribution and energy curves available in the Structure Assessment module of Swiss Model, and with the ProSA server [24]. To check the stability, the structures were subjected to molecular dynamics (MD) simulations of 100 ns (ns). Both structures were previously minimized and equilibrated with NVT/NPT phases, using GROMACS v5.1 [25]. The Amber99SB-ILDN protein force-field [26], a modified Berendsen thermostat [27], and the Parrinello-Rahman barostat [28] were used during the MD phases. The complex was solvated in a cubic box of water with periodic boundary conditions at a distance of at least 10 Å. Counterions were included in the solvent to make the box neutral. The Particle Mesh Ewald (PME) method was used to calculate the electrostatic interactions with 1.0 nm short-range threshold [29]. We used a timestep of 2 fs (fs)

and a temperature of 310 K. RMSD curves were calculated to monitor the protein flexibility during the simulated time.

3. Results

3.1. The sequence of Leishmania NMNATs contains specific insertions

From the multiple alignments between the amino acid sequences of NMNAT belonging to the five *Leishmania* species and the human isoenzymes, we identified a set of regions present exclusively in the parasite (Figure 1). These regions are insertions that in particular are structural components of the *Leishmania* genus. In addition, potential phosphorylation sites were found in the insertions, a post-translational modification that has been found in NMNATs from organisms such as *Trypanosoma cruzi* [30] and NMNAT-1 in humans [31].

The recombinant protein $6\text{His}\Delta_{241-249}\text{Lb}\text{NMNAT}$ was expressed in the bacterial *E. coli* BL21 (DE3) system as insoluble aggregates (Figure 2A). Despite modifications in expression conditions, including induction time, and co-expression with chaperones and foldases (DnaK, DnaE, GrpE, GroES, and GroEL), the protein was immunodetected in inclusion bodies (Figure 2B). As an alternative strategy the *E. coli* SHuffle strain was used, which partially favored the correct folding of the recombinant, allowing its expression in a soluble manner. The $6\text{His}-\Delta_{241}$. $_{249}\text{Lb}\text{NMNAT}$ and 6His-LbNMNAT proteins were partially purified by nickel affinity chromatography (Figure 3A) and immunodetected due to the presence of the histidine tag (Figure 3B).

3.2. Insertion 241–249 is indispensable for the enzymatic activity of LbNMNAT

In order to evaluate the catalytic activity of the mutated recombinant, we performed direct enzymatic assays, where the partially purified protein was incubated with the substrates (ATP, NMN and its cofactor Mg^{2+}) to follow the production of NAD⁺ by RP-HPLC (Figure 4A-C). These results were contrasted with coupled enzymatic assays measured spectrophotometrically at 340 nm, where the NAD⁺ product of the NMNAT synthesis is reduced to NADH by the action of the enzyme alcohol dehydrogenase (Figure 4D). By both methods, it was observed that the 6His- $\Delta_{241-249}$ LbNMNAT protein does not report adenylyl-transferase activity, which evidences how the exclusive insertion of the C-terminal region of the LbNMNAT is necessary for its *in vitro* enzymatic function.

Regarding the models and simulations, we generated a version of the LbNMNAT structure with the insertion (a red motif in Figure 5A) and without it (Figure 5B). The $\Delta_{241-249}$ LbNMNAT protein-NAD binding sites were predicted by COFACTOR [22] and COACH [32] tools based on the structure obtained with I-TASSER (Figure 5C). The result suggests a similar localization of NAD for the predicted binding site compared to HsNMNAT-1 co-crystallized with NAD (PDB: 1KQN), sharing similar interactions with residues His59 (electrostatic, Pi-Anion) and Thr130 (hydrogen bond) present in the catalytic site. However, the adenosine 5'-phosphate structure in NAD is sterically hindered by residues Asp172, Leu173, and Arg202 which may cause effects in the stabilization of the

		Р	20	Р	Р	40		60			80			100
LbNMNAT	MLSSTAAPYAL	RTDKLK	PLEGCAA	ASCPTSTA	EAASOVT	PLLOPV	VLVICGSFN	PIHNAHLKI	LYDAAKRS	IEGVDGR	VVLGGFLS	SPVGDA	YGKPGLR	CAAD 100
LdNMNAT	MPPSTAAPYAL	RTDKLK	PLEGCVA	AGCSVSPA	EAASOVA	SSLOPY	VLAICGSEN	PIHNAHLKI	LYDAAKRS	IDGAHGR	VVLGGFL	SPVGDA	YRKPGLR	SAAD 100
LINMNAT	MPPSTAAPYAL	RTDKLK	PLEGCV	AGCSVSPA	FAASOVA	SSLOPV	VLAICGSEN	PIHNAHLKI	YDAAKRS	IDGAHGR	VVLGGEL	SPVGDA	YRKPGLR	SAAD 100
LmNMNAT	MPPSTAAPYAL	RTDKLK	PLEGYV	AGCSVSPA	EVASOVA	SSLOPY	VLAICGSEN	PIHNAHLKI	LYDAAKRS	VDGADGR	VVLGGFL	SPVGDA	YGKPGLR	SAAD 100
LmxNMNAT	MPPSTAAPYAL	OTDKIK	PLEGCV	AGCSASTA	EMASOGA	SSLOPV	VLAICGSEN	PIHNAHLKI	VDAAKOS	IDGADGH	VVLGGFL	SPVGDA	YRKPGLH	SAAD 100
HsNMNAT-1	MENSEKTE					V	VILLACGSEN	PITNMHLRI	FELAKDY	MNGTGRY	TVVKGII	SPVGDA	YKKKGLI	PAYH 65
HsNMNAT-2	MTETTKTH					V	LLACGSEN	PITKGHION	VEERARDY	LHKTGRF	IVIGGIVS	SPVHDS	YGKOGLV	SSRH 65
HsNMNAT-3	MKSRIP					V	VLLACGSEN	PITNMHLRM	MEEVARDH	LHOTGMY	OVIOGIIS	SPVNDT	YGKKDLA	ASHH 63
Cancernation	MKJKIT						VEEACUSIN				QVIQUII		IUKKDEA	-
Conservation														
			120			140		160			180			200
LbNMNAT	RVOVMEKALCH	HPELNV	отжесос	OPTYTRTF	FVLRALE	EHVNAW	Y A						0	SEPA 151
LdNMNAT	RFOIMRKALCH	HPDLNV	DTWECOO	OPVYTRTE	FVLOALE	EHVNAW	Y A						E	SEPA 151
LINMNAT	RLOIMRKALCH	HPDLNV	DTWECOO	OPVYTRTF	FVLOALE	EHVNAW	Y A						E	SEPA 151
LmNMNAT	RVHIMRKALCH	HPDLNV	DTWECOO	OPVYMRTE	FVLOALE	EHINAW	Y A						E	SEPA 151
LmxNMNAT	RVOIMRKALCH	HPDLNV	DTWECO	OPVYTRTE	FVLOALE	EHVNAW	Y A						0	SEPA 151
HsNMNAT-1	RVIMAELATKN	SKWVEV	DTWESLO	OKEWKETL	KVLRHHO	EKLEA -	S D C D	HOONSPILI	ERPGRKRK	WTETODS	\$0		KKSLE	РКТК 146
HsNMNAT-2	RLIMCOLAVON	SDWIRV	DPWECY	ODTWOTTC	SVLEHHR	DLMKRV	TGCILSNVN	TPSMTPVI	GOPONETP	OPIYONS	NVATKPT/	AAKILG	KVGESLS	RICC 165
HsNMNAT-3	RVAMARLALOT	SDWIRV	DPWESE	OAOWMETV	KVLRHHH	SKLLR -	S P	POMI	EGP		DH		GKALF	STPA 125
Concernation				2							511		UNITED	
Conservation														
			220			240 		260 			280 I			300 I
LbNMNAT	AMEWLTS		220 I	HGRHVRVV	FACGADL	240 I F F S FWR	PGCWSLCLL	260 R Q L L D S F P I	LVVVQRGG	ARGNISD	280 I S D D F A R V (CQTAPL	LWETAED	300 I GERI 235
LbNMNAT LdNMNAT	AMEWL T S AMEWL T S		220 	H G R H V R V V H D R R V R V V	F A C G A D L F V C G A D L	240 F F S FWR F S S FW I	P G C W S L C L L P G C W Q L R L L	260 RQLLDSFP RQLLDSFH	L V V V Q R G G L V V V H R E G	A R G N I S D A R G C V R G	280 I S D D F A R V (A D D F A H I (CQTAPL CQTAPL	LWETAED LSETAVD	300 I GERI 235 GEKI 235
LbNMNAT LdNMNAT LiNMNAT	AMEWLTS AMEWLTS AMKWLTS		220 I 	H G R H V R V V H D R R V R V V H D R R V R V V	F A C G A D L F V C G A D L F V C G A D L	240 I F F S FWR F S S FWI F S S FWI	P G C W S L C L L P G C W Q L R L L P G C W Q L R L L	260 R Q L L D S F P I R Q L L D S F H I R Q L L D S F H I	L V V V Q R G G L V V V H R E G L V V V H R E G	A R G N I S D A R G C V R G A R G C V R G	280 I S D D F A R V (A D D F A H I (A D D F A H I (CQTAPL CQTAPL CQTAPL	LWE T A E D L S E T A V D L S E T A V D	300 G E R I 235 G E K I 235 G E K I 235 G E K I 235
LbNMNAT LdNMNAT LiNMNAT LmNMNAT	AMEWLTS AMEWLTS AMKWLTS AVEWLAS		220 I 	H G R H V R V V H D R R V R V V H D R R V R V V H D R R V R V A	F A C G A D L F V C G A D L F V C G A D L F V C G A D L	240 I F F S FWR F S S FWI F S S FWI F S S FWI	P G C W S L C L L P G C W Q L R L L P G C W Q L R L L P G C W Q L R L L	260 I R Q L L D S F P I R Q L L D S F H I R Q L L D S F H I R Q L L D S F H I	L V V V Q R G G L V V V H R E G L V V V H R E G L V V V H R D G L V V V H R D G	A R G N I S D A R G C V R G A R G C V R G G R G C V R G	280 I S D D F A R V (A D D F A H I (A D D F A H I (A D D F A H V (CQTAPL CQTAPL CQTAPL CQTAPL CQTAPL	LWETAED LSETAVD LSETAVD LSETAEG	300 G E R I 235 G E K I 235 G E K I 235 G E K I 235 G E K I 235
LbNMNAT LdNMNAT LiNMNAT LmNMNAT LmXNMNAT	AMEWLTS AMEWLTS AMKWLTS A VEWLAS EMEWLAS		220 I 	H G R H V R V V H D R R V R V V H D R R V R V V H D R R V R V A R D L R V R V V	F A C G A D L F V C G A D L	240 I F F S FWR F S S FWI F S S FWI F S S FWI F S S FWI	P G C W S L C L L P G C W Q L R L L P G C W Q L R L L P G C W Q L R L L P G C W P L R L L	260 I R Q L L D S F P I R Q L L D S F H I R Q L L D S F H I R Q L L D S F H I R R L L D S F R V	L V V V Q R G G L V V V H R E G L V V V H R E G L V V V H R D G V V V H R D G	A R G N I S D A R G C V R G A R G C V R G G R G C V R G A R G G V R G	280 I S D D F A R V (A D D F A H I (A D D F A H I (A D D F A H V (A D D F A H V (CQTAPL CQTAPL CQTAPL CQTAPL CQTAPL CQTAPL	LWETAED LSETAVD LSETAVD LSETAEG LSETAEG	300 G E R I 235 G E K I 235
LbNMNAT LdNMNAT LiNMNAT LmNMNAT LmxNMNAT HsNMNAT-1	AMEWLTS AMEWLTS AMKWLTS AVEWLAS EMEWLAS AVPKVK		220 I 	H G R H V R V V H D R R V R V V H D R R V R V V H D R R V R V A R D L R V R V V	F A C G A D L F V C G A D L L L C G A D L	240 I F F S FWR F S S FWI F S S FWI F S S FWI F S S FWI L E S F A V	P G C W S L C L L P G C W Q L R L L P G C W Q L R L L P G C W Q L R L L P G C W P L R L L P N L W K S E D I	260 R Q L L D S F P I R Q L L D S F H I R Q L L D S F H I R Q L L D S F H I R R L L D S F R V T Q I V A N Y G I	L V V V Q R G G L V V V H R E G L V V V H R E G L V V V H R D G V V V V H R D G L I C V T R A G	A R G N I S D A R G C V R G A R G C V R G G R G C V R G A R G G V R G N	280 I S D D F A R V (A D D F A H I (A D D F A H I (A D D F A H V (A D D F A H V (D A Q K F	CQTAPL CQTAPL CQTAPL CQTAPL CQTAPL IYESDV	LWETAED LSETAVD LSETAVD LSETAEG LSETAEG LSETAEG LWKHRSN	300 G E R I 235 G E K I 235 I H V V 213
LbNMNAT LdNMNAT LiNMNAT LmNMNAT LmxNMNAT HsNMNAT-1 HsNMNAT-2	AMEWLTS AMEWLTS AMKWLTS AVEWLAS EMEWLAS AVPKVK AVPKVK VRPPVERFTFV	DENANL	220 I 	H G R H V R V V H D R R V R V V H D R R V R V V H D R R V R V A R D L R V R V V E E I E L R I L	F A C G A D L F V C G A D L L L C G A D L L L C G S D L	240 I F F S FWR F S S FWI F S S FWI F S S FWI L E S F A V L E S F C I	P G CWS L C L L P G CWQ L R L L P G CWQ L R L L P G CWQ L R L L P G CWP L R L L P G CWP L R L L P N LWK S E D I P G LWN E A DM	260 I R Q L L D S F P I R Q L L D S F H I R Q L L D S F H I R Q L L D S F H I R R L L D S F R V T Q I V A N Y G I E V I V G D F G	L V V V Q R G G L V V V H R E G L V V V H R E G L V V V H R D G V V V V H R D G L I C V T R A G I V V V P R D A	A R G N I S D A R G C V R G A R G C V R G G R G C V R G A R G G V R G N	280 I S D D F A R V (A D D F A H I (A D D F A H I (A D D F A H V (A D D F A H V (- D A Q K F I - D T D R I M	CQTAPL CQTAPL CQTAPL CQTAPL CQTAPL IYESDV MNHSSI	LWETAED LSETAVD LSETAVD LSETAEG LSETAEG LWKHRSN LRKYKNN	300 G E R I 235 G E K I 235 G E K I 235 G E K I 235 G E K I 235 I H V V 213 I M V V 257
LbNMNAT LdNMNAT LiNMNAT LmNMNAT HsNMNAT-1 HsNMNAT-2 HsNMNAT-3	AMEWLTS AMEWLTS AMKWLTS AVEWLAS EMEWLAS AVPKVK VRPVERFTFV AVPELK	DENANL	220 I 	H G R H V R V V H D R R V R V V H D R R V R V V H D R R V R V X A D L R V R V V E E I E L R I L	F A C G A D L F V C G A D L L L C G A D L L L C G S D L L L C G A D V	240 1 F F S FWR F S S FWI F S S FWI F S S FWI L E S F A V L E S F C I L KT F Q T	P G CWS L C L L P G CWQ L R L L P G CWQ L R L L P G CWQ L R L L P G CWP L R L L P N LWK S E D I P N LWK D A H I	260 I R Q L L D S F P I R Q L L D S F H I R Q L L D S F H I R R L L D S F R V T Q I V A N Y G I E V I V G D F G Q E I V E K F G I	L V V V Q R G G L V V V H R E G L V V V H R E G V V V H R D G V V V V H R D G L I C V T R A G I V V V P R D A L V C V G R V G	A R G N I S D A R G C V R G A R G C V R G G R G C V R G A R G G V R G N A H	280 I S D D F A R V (A D D F A H I (A D D F A H I (A D D F A H V (A D D F A H V (D A Q K F D T D R I N D P K G Y	CQTAPL CQTAPL CQTAPL CQTAPL CQTAPL IYESDV MNHSSI IAESPI	LWETAED LSETAVD LSETAVD LSETAEG LSETAEG LWKHRSN LRKYKNN LRKYKNN	300 J G E R I 235 G E K I 235 G E K I 235 G E K I 235 G E K I 235 J H V V 213 I M V V 257 I H L A 192
LbNMNAT LdNMNAT LiNMNAT LmXNMNAT HSNMNAT-1 HSNMNAT-2 HSNMNAT-3 Conservation	AMEWLTS AMEWLTS AMKWLTS AVEWLAS EMEWLAS AVPKVK VRPVERFTFV AVPELK	DENANL	220 I I I G T V M R Y I	H G R H V R V V H D R R V R V V H D R R V R V R V H D R R V R V R V R D L R V R V R E E I E L R I L	F A C G A D L F V C G A D L F V C G A D L F V C G A D L L C G A D L L L C G A D L L L C G A D V	240 1 F F S FWR F S S FWI F S S FWI F S S FWI L E S F A V L E S F C I L KT F Q T	P G C W S L C L L P G C W Q L R L L P G C W Q L R L L P G C W Q L R L L P G C W P L R L L P N L W K S E D I P G L W N E A D M P N L W K D A H I	260 I R Q L L D S F H I R Q L L D S F H I R Q L L D S F H I R R L L D S F R I R R L L D S F R I C I V A N Y G I E V I V G D F G Q E I V E K F G I	L V V V Q R G G L V V V H R E G L V V V H R D G V V V V H R D G L I C V T R A G I V V V P R D A L V C V G R V G	A R G N I S D A R G C V R G G R G C V R G G R G C V R G A R G G V R G N A H	280 I S D D F A R V (A D D F A H I (A D D F A H V (A D D F A H V (A D D F A H V (D A Q K F D T D R I N D P K G Y	CQTAPL CQTAPL CQTAPL CQTAPL CQTAPL IYESDV MNHSSI IAESPI	LWETAED LSETAVD LSETAVG LSETAEG LSETAEG LWKHRSN LRKYKNN LRKYKNN	300 I G E R I 235 G E K I 235 G E K I 235 G E K I 235 G E K I 235 I H V V 213 I M V V 257 I H L A 192
LbNMNAT LdNMNAT LiNMNAT LmNMNAT HsNMNAT-1 HsNMNAT-2 HsNMNAT-3 Conservation	AMEWLTS AMEWLTS AMKWLTS AVEWLAS EMEWLAS AVPKVK VRPVERFTFV AVPELK	DENANL	220 1 	H G R H V R V V H D R R V R V V H D R R V R V V H D R R V R V V A D L R V R V V E E I E L R I L	F A C G A D L F V C G A D L F V C G A D L F V C G A D L L C G A D L L L C G A D L L L C G A D L	240 I F F S FWR F S S FWI F S S FWI F S S FWI L E S F A V L E S F A V L E S F C I L K T F Q T	P G CWS L C L L P G CWQ L R L L P N LWK S E D I P N LWK D A H I	260 I RQLLDSFPI RQLLDSFFI RQLLDSFFI RQLLDSFFI RQLLDSFFF RRLLDSFR TQIVANYG EVIVGDFG QEIVEKFG 360	L V V V Q R G G L V V V H R E G L V V V H R D G V V V V H R D G L I C V T R A G L I C V T R A G L V C V G R V G	A R G N I S D A R G C V R G G R G C V R G G R G C V R G A R G G V R G N A H	280 I S D D F A R V (A D D F A H I (A D D F A H V (A D D F A H V (A D D F A H V (- D A Q K F - D T D R I N - D P K G Y 380	C Q T A P L C Q T A P L I Y E S D V MN H S S I I A E S P I	LWETAED LSETAVD LSETAEG LSETAEG LWKHRSN LRKYKNN LRKYKNN	300 G E R I 235 G E K I 235 G E K I 235 G E K I 235 I H V V 213 I M V V 213 I M V V 217 I H L A 192
LbNMNAT LdNMNAT LmNMNAT LmXNMNAT HsNMNAT-1 HsNMNAT-3 Conservation	AMEWLTS AMEWLTS AVEWLAS AVEWLAS AVPKVK VRPVERFTFV AVPELK	DENANL	220 1 	H G R H V R V V H D R R V R V V H D R R V R V V H D R R V R V V E D L R V R V V E E I E L R I L	F A C G A D L F V C G A D L F V C G A D L F V C G A D L L L C G A D L L L C G A D V	240 I F F S FWR F S S FWI F S S FWI F S S FWI L E S F A V L E S F A V L E S F C I L K T F Q T 340 I	P G CWS L C L L P G CWQ L R L L P G CWQ L R L L P G CWQ L R L L P G CWP L R L L P N LWK SE D I P G LWN E A DM P N LWK DA H I	260 I RQLLDSFPI RQLLDSFHI RQLLDSFHI RQLLDSFRT RQLLDSFRT TQIVANYGI EVIVGDFG QEIVEKFGI 360	L V V V Q R G G L V V V H R E G L V V V H R E G L V V V H R D G V V V V H R D G L I C V T R A G I V V V P R D A L V C V G R V G	A R G N I S D A R G C V R G A R G C V R G G R G C V R G A R G G V R G N A H	280 I S D D F A R V (A D D F A H I (A D D F A H V (CQTAPL CQTAPL CQTAPL CQTAPL CQTAPL IYESDV MNHSSI IAESPI	LWETAED LSETAVD LSETAVD LSETAEG LSETAEG LWKHRSN LRKYKNN LRMHQHN	300 G E R I 235 G E K I 235 G E K I 235 G E K I 235 I H V V 213 I M V V 213 I M V V 217 I H L A 192
LbNMNAT LdNMNAT LiNMNAT LmxNMNAT HSNMNAT-1 HSNMNAT-3 Conservation	AMEWLTS AMEWLTS AVEWLAS EMEWLAS VRPPVERFTFV AVPELK P P EIDMLRYTFTF	DENANL P AAFSVP	220 I 	H G R H V R V V H D R R V R V V H D R R V R V V H D R R V R V A R D L R V R V A E E I E L R I L E E I E L R I L	F A C G A D L F V C G A D L F V C G A D L F V C G A D L L C G A D L L L C G A D V L A KMP V D L	240 I F F S F WR F S S FWI F S S FWI F S S FWI F S S FWI L E S F A V L E S F A V L E S F C I L K T F Q T 340 I E V T A R T	P G CWS L C L L P G CWQ L R L L P G CWQ L R L L P G CWQ L R L L P G CWP L R L L P M LWK S E D I P G LWN E A DM P N LWK D A H I	260 I RQLLDSFPI RQLLDSFPI RQLLDSFH RQLLDSFH TQIVANYG TQIVANYG EVIVGDFG QEIVEKFGI 300 I LPEAAIQL	LVVVQRGG LVVVHREG LVVVHRDG LVVVHRDG LICVTRAG IVVVPRDA LVCVGRVG	A R G N I S D A R G C V R G A R G C V R G G R G C V R G A R G C V R G N A H	280 I S D D F A R V (A D D F A H I (A D D F A H I (A D D F A H V (- D A Q K F - D T D R I N - D P K G Y - 380 I	CQTAPL CQTAPL CQTAPL CQTAPL IYESDV MNHSSI IAESPI	LWETAED LSETAVD LSETAVD LSETAEG LSETAEG LWKHRSN LRKYKNN LRMHQHN EGLPS 307	300 G E R I 235 G E K I 235 G E K I 235 G E K I 235 G E K I 235 I H V V 213 I M V V 257 I H L A 192
LbNMNAT LdNMNAT LiNMNAT LmxNMNAT HSNMNAT-1 HSNMNAT-2 HSNMNAT-3 Conservation LbNMNAT	AM EWL T S AM EWL T S AW EWL A S EM EWL A S V P F V E R F T F V A V P E L K P P E I DML R Y T F T F E I DMS R Y T F T F	DENANL P AAFSVP ATFSAP	220 I 	H G R H V R V V H D R R V R V V H D R R V R V V H D R R V R V A R D L R V R V V E E I E L R I L E E I E L R I L	F A C G A D L F V C G A D L F V C G A D L F V C G A D L L C G A D V L A KM P V D L A KM P V D	240 I F F S FWR F S S FWI F S S FWI F S S FWI L E S F A V L E S F A V A V T A R T.	P G CWS L C L L P G CWQ L R L L P G CWQ L R L L P G CWQ L R L L P G CWP L R L L P N LWK S E D I P G LWN E A DM P N LWK D A H I	260 I RQLLDSFPI RQLLDSFPI RQLLDSFFI RQLLDSFFI RQLLDSFFI RQLLDSFFI RQLVSFFI EVIVGDFG QEIVEKFGI 360 I LPEAAIQLV LPEAAVQS	LVVVQRGG LVVVHREG LVVVHRDG VVVVHRDG IVVVPRDA LVCVGRVG VVVPRDA LVCVGRVG	A R G N I S D A R G C V R G A R G C V R G G R G C V R G A R G C V R G A R G V R G H	280 I S D D F A R V (A D D F A H I (A D D F A H V (CQTAPL CQTAPL CQTAPL CQTAPL IYESDV MNHSSI IAESPI	LWETAED LSETAVD LSETAEG LSETAEG LSETAEG LWKHRSN LRKYKNN LRKYKNN LRMHQHN EGLPS 307 EVLTS 307	300 G E R I 235 G E K I 235 I H V V 213 I M V V 257 I H L A 192
LbNMNAT LdNMNAT LiNMNAT LmNMNAT HsNMNAT-1 HsNMNAT-2 HsNMNAT-3 Conservation LbNMNAT LdNMNAT LiNMNAT	AM EWL T S AM EWL T S AM KWL T S AV EWL AS EM EWL AS AV PK VK V R P V E R FT FV A V P E L K P P E 1 DM L RYT FT F E 1 DM S RYT FT F E 1 DM S RYT FT F	P A A F S V P A T F S A P A T F S A P	220 1 	H G R H V R V V H D R R V R V V H D R R V R V V H D R V R V V V E E I E L R I L E E I E L R I L A V R N A V M E A V R A A A ME A V R A A A ME	F A C G A D L F V C G A D L F V C G A D L F V C G A D L L L C G A D L L L C G S D L L L C G A D V L A KM P V D L A KM P V D L A KT P L G	240 1 F S FWR F S S FWI F S S FWI F S S FWI L E S F AV L E S F AV L E S F AV L E S F AV E S F AV E V T A R T. D A V T R A.	P G CWS L C L L P G CWQ L R L K P N LWK S E D I P N LWK D A H I A L Q S R L R T M A L Q S R L R AM	260 1 RQLLDSFPI RQLLDSFFH RQLLDSFHI RQLLDSFHI RQLLDSFHI QUIVSFHI EVIVGDFG QUIVEKFGI 360 1 LPEAAIQLY LPEAAVQSY LPEAAVQSY	LVVVQRGG LVVVHRG LVVVHRG LVVVHRDG VVVVHRDG IVVVPRDA LVCVGRVG VVD-LY- VVC-LY- VVE-LY-	A R G N I S D A R G C V R G A R G C V R G G R G C V R G N H	280 I S D D F A R V (A D D F A H I (A D D F A H V (CQTAPL CQTAPL CQTAPL CQTAPL IYESDV MNHSSI IAESPI	LWETAED LSETAVD LSETAEG LSETAEG LWKHRSN LRKYKNN LRKYKNN LRMHQHN EGLPS 307 EVLTS 307	300 G E R I 235 G E K I 235 I H V V 213 I H V V 213 I H V A 192
LbNMNAT LdNMNAT LmNNNAT LmXNMNAT HSNMNAT-1 HSNMNAT-2 HSNMNAT-3 Conservation LbNMNAT LdNMNAT LiNMNAT	AMEWLTS AMEWLTS AVEWLAS AVEWLAS AVPKVK VRPVERFTFV AVPELK P P E I DMLRYTFTF E I DMSRYTFTF E I DMSRYTFTF	P A A F S V P A T F S A P A T F S A P	220 1 	H G R H V R V V H D R R V R V V H D R R V R V V H D R R V R V V E E I E L R I L E E I E L R I L A V R A A ME A V R A A ME A V R A A ME	FACGADL FVCGADL FVCGADL LCGADL LLCGADL LLCGADL LLCGADV LAKMPVD LAKMPVD LAKTPLG LAKTPLG	240 1 F S FWR F S SFWI F S SFWI F S SFWI L E S FAV L E S FAV L E S FAV L E S FAV L E S TAT AU AU AU AU AU AU AU AU AU AU	P G CWS L C L L P G CWQ L R L L P G CWQ L R L L P G CWQ L R L L P G CWP L R L L P N LWK SE D I P G LWN E A D M P N LWK D A H I A L Q S R L R A M A L Q S R L R A M	260 I RQLLDSFPI RQLLDSFHI RQLLDSFHI RQLLDSFRT TQIVANYGI EVIVGDFG QEIVEKFGI 360 LPEAAVQS LPEAAVQS LPEAAVQS LPEAAVQS	LVVVQRGG LVVVHRG LVVVHRG LVVVHRDG LVVVHRDG LICVTRAG IVVVPRDA LVCVGRVG VVD - LY- VVD - LY- VVE - LY-	A R G N I S D A R G C V R G G R G C V R G (G R G C V R G N A H	280 I S D D F A H V A D D F A H I (A D D F A H V A D D F A H V - D A Q K F - D T D R I N - D P K G Y 3800 I	CQTAPL CQTAPL CQTAPL CQTAPL IYESDV MNHSSI IAESPI	LWETAED LSETAVD LSETAEG LSETAEG LWKHRSN LRKYKNN LRKYKNN EGLPS 307 EVLTS 307	300 G E R I 235 G E K I 235 I H V V 213 I M V V 213 I H L A 192
LbNMNAT LdNMNAT LiNMNAT LmXNMNAT HSNMNAT-1 HSNMNAT-3 Conservation LbNMNAT LdNMNAT LiNMNAT LiNMNAT	AMEWLTS AMEWLTS AVEWLAS EMEWLAS VRPPVERFTFV AVPELK P P E 1 DMLRYTFTF E 1 DMSRYTFTF E 1 DMSRYTFTF E 1 DMSRYTFTF	P AAFSVP ATFSAP ATFSAP ATFSAP ATFSAP	220 1 320 G T V M R Y I D D T S S T / D D T S S T /	H G R H V R V V H D R R V R V V H D R R V R V V H D R R V R V V G L R V R V V E E I E L R I L E E I E L R I L A V R A A ME A V R A A AME A V R A A AME A V R A A AME	FACGADL FVCGADL FVCGADL LCGADL LLCGADL LLCGADL LLCGADV LAKMPVD LAKTPLG LAKTPLG LAKTPLS	240 F S FWR F S S FWI F S S FWI F S S FWI L E S F A V L E S F A V A V A A V L E S F A V A V A A V L E S F A V A V A A V A A V A A V A A A V A A A A V A	P G CWS L C L L P G CWQ L R L L P G CWQ L R L L P G CWQ L R L L P G CWP L R L L P M LWK S E D I P G CWP L R L A L M L Q S R L R T M A L Q S R L R A M A L Q S R L R A M A L Q S R L R A M	260 1 RQLLDSFP1 RQLLDSFP1 RQLLDSFP1 RQLLDSFP1 TQIVANYGI EVIVGDFG QEIVEKFG 1 LPEAAVQSV LPEAAVQSV LPEAAVQSV	LVVVQRGG LVVVHREG LVVVHREG LVVVHRDG LICVTRAG IVVVPRDA LVCVGRVG VVD - LY- VVE - LY- VVE - LY- VVE - LY-	A R G N I S D A R G C V R G A R G C V R G G R G C V R G N H	280 I S D D F A R V (A D D F A H I (A D D F A H I (A D D F A H V (- D A Q K F - D T D R I N - D P K G Y - 1 - 0 P K G Y	CQTAPL CQTAPL CQTAPL CQTAPL IYESDV MNHSSI IAESPI G G G	LWETAED LSETAVD LSETAEG LSETAEG LWKHRSN LRKYKNN LRMHQHN EGLPS 307 EVLTS 307 EVLTS 307	300 G E R I 235 G E K I 235 I H VV 213 I M V V 257 I H L A 192
LbNMNAT LdNMNAT LiNMNAT LmxNMNAT HsNMNAT-1 HsNMNAT-3 Conservation LbNMNAT LiNMNAT LiNMNAT LmxNMNAT HsNMNAT-1	AMEWLTS AMEWLTS AVEWLAS EMEWLAS VRPVERFTFV AVPELK P P EIDMLRYTFTF EIDMSRYTFTF EIDMSRYTFTF EIDMSRYTFTF VIDMSRYTFTF	P AAFSVP ATFSAP ATFSAP ATFSAP ATFSAP ATFSAP ATFSAP	220 1 320 1 5 5 5 5 5 5 5 5 5 5 5 5 5	H G R H V R V V H D R R V R V V H D R R V R V V H D R R V R V A C L R V R V A E E I E L R I L A V R A A ME A V R A A A ME	F A C G A D L F V C G A D L F V C G A D L F V C G A D L L C G A D V L A KMP V D L A K T P L G L A K T P L G	240 1 F S F WR F S S FWI F S S FWI F S S FWI E S F WI L E S F AV L E S F AV L E S F AV L E S F AV L E S F AV AVTRA. D A V TRA. D A V TRA. D A V RA. RR	P G CWS L C L L P G CWQ L R L L P G CWQ L R L L P G CWQ L R L L P G CWP L R L L P N LWK S E D I P G CWN E A DM P N LWK D A H I A L Q S R L R A M A L Q S R L R A M A L Q S R L R A M A L Q S R L R A M	260 1 RQLLDSFPI RQLLDSFPI RQLLDSFPI RQLLDSFFI RQLLDSFFI TQIVANYGI EVIVGDFG QEIVEKFGI 1 LPEAAIQLS LPEAAVQSS LPEAAVQSS LPEAAVQSS VPDLVQEY	L V V V Q R G G L V V V H R E G L V V V H R E G L V V V H R D G L I C V T R A G I V V V P R D A L V C V G R V G V V C - L Y - V V E - L Y - V V E - L Y - V V E - L Y -	A R G N I S D A R G C V R G A R G C V R G G R G C V R G A R G G V R G V R H	280 I S D D F A R V (A D D F A H I (A D D F A H V (CQTAPL CQTAPL CQTAPL CQTAPL CQTAPL I YESDV MNHSSI I AESPI G G G LQRNTA	LWETAED LSETAVD LSETAEG LSETAEG LWKHRSN LRKYKNN LRKYKNN LRMHQHN EGLPS 300 EVLTS 300 EVLTS 300 EVLTS 300 EVLTS 300 EVLTS 300	300 G E R I 235 G E K I 235 I H V V 213 I M V V 257 I H L A 192
LbNMNAT LdNMNAT LiNMNAT LmxNMNAT HSNMNAT-1 HSNMNAT-2 HSNMNAT-3 Conservation LbNMNAT LdNMNAT LmNNAT LmNMNAT HSNMNAT-1 HSNMNAT-2	AM EWL T S AM EWL T S AM EWL T S AV EWL AS EM EWL AS V P P V E R FT FV AV P E L K P P E I DML R YT FT F E I DMS R YT FT F E I DMS R YT FT F E I DMS R YT FT F F I DMS R YT FT F F I DMS R YT FT F N EWI A	P A A F S V P A T F S A P A T L S A P A T L S A P A T L S A P	220 1 320 G T V M R Y I D D T S S T / D D T S S T / S /	H G R H V R V V H D R R V R V V H D R R V R V V H D R R V R V A R D L R V R V A E E I E L R I L A V R A A M E A V R A A A M E	F A C G A D L F V C G A D L F V C G A D L F V C G A D L L C G A D V L A KM P V D L A KT P L G L A KT P L G L A KT P L S L A KT P L S	240 1 F S F WR F S S FWI F S S FWI F S S FWI L S F AV L E S F AV L E S F AV L E S F AV 1 E V T AR T. D A V T R.A. D A V T R.A. C A V T R.A. D	P G CWS L C L L P G CWQ L R L L P N LWK S E D I P G LWN E A DM P N LWK D A H I A L Q S R L R A M A L Q S R L R A M A L Q S R L R A M G Q S I R Y L G D G H V V D Y L	260 1 RQLLDSFPI RQLLDSFPI RQLLDSFPI RQLLDSFFI RQLLDSFFI RQLLDSFFI QEIVEKFGI 200 200 200 200 200 200 200 20	L V V V Q R G G L V V V H R E G L V V V H R D G V V V V H R D G I V V V P R D A L V C V G R V G V V C - L Y - V V E - L Y - I E K H N L Y S I L K S Q L Y I	A R G N I S D A R G C V R G A R G C V R G A R G C V R G A R G G V R G A R G V R G V R G V R G A R G V	280 I S D D F A R V (A D D F A H I (A D D F A H V (CQTAPL CQTAPL CQTAPL CQTAPL I YESDV MNHSSI I AESPI G G G LQRNTA	LWETAED LSETAVD LSETAEG LSETAEG LSETAEG LWKHRSN LRKYKNN LRKYKNN LRMHQHN EGLPS 307 EVLTS 307 EVLTS 307 KVLTS 307 EVLTS 307 EVLTS 307	300 G E R I 235 G E K I 235 I H V V 213 I M V V 257 I H L A 192
LbNMNAT LdNMNAT LiNMNAT HSNMNAT-1 HSNMNAT-3 Conservation LbNMNAT LdNMNAT LiNMNAT LiNMNAT HSNMNAT-1 HSNMNAT-1 HSNMNAT-2 HSNMNAT-3	AMEWLTS AMEWLTS AVEWLAS AVEWLAS AVPKVK VRPPVERFTFV AVPELK F EIDMLRYTFTF EIDMLRYTFTF EIDMSRYTFTF EIDMSRYTFTF F IDMSRYTFTF NEWIA KDDIN	P A A F S V P A T F S A P A T L S A P - ND - - ND - ND - ND - NE -	220 1 	H G R H V R V V H D R R V R V V H D R R V R V V H D R R V R V V E E I E L R I L E E I E L R I L A V R A A ME A V R A A L - K S R L A L - Y I R R A L -	F A C G A D L F V C G A D L F V C G A D L F V C G A D L L C G A D L L C G A D L L C G A D V L A KMP V D L A KMP V D L A KT P L G L A KT P L G L A KT P L G	240 1 F S FWR F S S FWI F S S FWI F S S FWI L E S F A V L E S F A V A V T A A D A V T A A D A V T R A D D A V T A A D D A V T A A D D A V	P G CWS L C L L P G CWQ L R L L P G CWQ L R L L P G CWQ L R L L P G CWP L R L L P N LWK SE D I P G CWP L R L A D N LWK D A H I I I I I I I I I I I I I I I I I I I	260 1 RQLLDSFPI RQLLDSFFH RQLLDSFHI RQLLDSFHI RQLLDSFHI QEIVEKFGI 200 200 200 200 200 200 200 20	L V V V Q R G G L V V V H R E G L V V V H R D G V V V V H R D G I V V V P D A L V C V G R V G V V C - L Y - V V E - L Y - V V E - L Y - V V E - L Y - I E K H N L Y S I L K S Q L Y I I K D H G L Y T	A R G N I S D A R G C V R G G R G C V R G V R G C V R G A R G C V R G N	280 I S D D F A R V (A D D F A H I (A D D F A H V (- D A Q K F - D A Q K F - D A Q K F - D T D R I N - D P K G Y 	C Q T A P L C Q T A P L I A E S P I I A E	LWE T A E D LS E T A V D LS E T A E G LS E T A E G LWKHR S N LR KY KN N LR KY KN N LR MH QH N E G L P S 307 E V L T S 307 E V L	300 G E R I 235 G E K I 235 I H V V 213 I H V V 213 I H V V 213 I H V A 192

Figure 1. Identification of specific regions of the NMNAT of *Leishmania* species. Multiple alignment of NMNATs from different *Leishmania* species and the human isoenzymes HsNMNAT1-3. (LbNMNAT: *L. braziliensis* MHOM/BR/75/M2904, Ld: *L. donovani*, Li: *L. infantum JPCM5*, Lm: *L. major Friendlin*, Lmx: *L. mexicana*). The green boxes indicate the specific regions whose positions correspond to the amino acids: 8–43 and 241–249. The yellow boxes indicate the two reported ATP-binding motifs: [H/T]xxH and SxxxxR [29]. Red P's indicate possible residues subject to phosphorylation. The identity between the sequences is proportional to the height of the green bars.



Figure 2. Expression of recombinant 6His- $\Delta_{241.249}$ LbNMNAT protein under different experimental conditions. (A). 10% SDS-PAGE, the soluble and insoluble fractions of different expression systems untransformed (-) or transformed (+) with plasmid pQE30- $\Delta_{241.249}$ LbNMNAT and induced with IPTG at 24 °C for different times were analyzed. Expression systems used: codon plus (CP), and co-expression with chaperones (Ch). Red arrows indicate the accumulation of 6His- $\Delta_{241.249}$ LbNMNAT protein in the insoluble fractions of the different expression systems. Proteins were visualized with Coomassie blue. (B). Additionally, samples were analyzed by Western blot using the primary anti-6His antibody and peroxidase development system. M, molecular weight marker in kDa (Fig. S1).



Figure 3. Purification and immunodetection of mutated and full-length proteins. (A) 12% SDS-PAGE. Coomassie R-250 staining. Lanes 1–2: partially purified recombinant proteins $6His-\Delta_{241-249}$ LbNMNAT (36 kDa) and 6His-LbNMNAT (37 kDa), respectively. (B) Western blot analysis using the primary anti-6His antibody and alkaline phosphatase development system. M: molecular weight marker in kDa (Fig. S2).

protein-ligand interaction. Also, wild-type and mutated protein models were subjected to MD simulations of 100 ns to observe the general atom fluctuations based on the evolution of the protein RMSD. We found that removing the connector influences the protein flexibility, making it more prone to be destabilized through time (Figure 5D).

4. Discussion

The experimental and computational characterization of the exclusive insertion in LbNMNAT have allowed us to analyze the important role of this connector for the enzyme stability and activity, becoming a



Figure 4. Removal of residues 241–249 of LbNMNAT affects its adenylyltransferase activity. NAD⁺ synthesis was verified by RP-HPLC. Corresponding peaks are indicated for the substrates NMN and ATP, as well as for the product NAD⁺. (A) Reaction buffer (negative activity control). (B) Wild type 6His-LbNMNAT protein. (C) $6His-\Delta_{241-249}LbNMNAT$ protein. (D) Coupled enzyme assays for recombinants 6His-LbNMNAT and $6His-\Delta_{241-249}LbNMNAT$.

relevant factor in the design of potential therapeutic alternatives. From an *in silico* perspective, the integral analysis through the predictive modeling of the LbNMNAT protein and the multiple alignment between the NMNATs sequences of different species was crucial to find an exclusive and conserved region at the carboxyl terminus of the parasite protein (residues 241–249). Moreover, it complements a previous investigation reporting insertions essential for its *in vitro* activity [18]. These regions are structural components particular to the parasite genus and located in potential sites of functional blockade.

The identification of unique insertions in NMNATs from other organisms has already been studied. Analyzes performed by Lau, C. et al. [35], have shown that the three genes for human NMNATs contain exons that encode isoforms with unique insertions of specific sequences, called ISTIDs (Isoform Specific Targeting and Interaction Domains), which are part of regions associated with regulatory functions and post-translational modifications [35]. The identification of ISTIDs as functional units in human NMNATs provides a molecular basis for studying important non-catalytic functions. Therefore, the study of deletional mutants of the NMNAT of *L. braziliensis* becomes a prospective tool to determine the role of specific regions for the enzyme's function.

In vitro evaluation of the enzymatic activity of the recombinant 6His- $\Delta_{241-249}$ LbNMNAT showed that the removal of residues 241–249, close to the ATP-binding motif at the C-terminus, affects its adenylyltransferase activity. When comparing the structures resolved by X-ray crystallography of different NMNATs, the presence of a structural element composed of a loop or a loop/ β -sheet combination is crucial to connect the Rossmann topology with the C-terminal domain [31]. Such a structural element is susceptible to conformational changes required for catalysis, and also provides a key contribution to the stabilization of the oligomeric assembly [31]. In the 6His- $\Delta_{241-249}$ LbNMNAT protein, the structure and composition of the β sheet that connects the Rossmann domain and the C-terminal domain is altered (Figure 5A), affecting the enzymatic activity despite the deletion is not directly involved or located nearby the catalytic residues. In this way, we demonstrate the structural and functional relevance of this region in the catalytic activity, which is necessary to evaluate its role in oligomerization, sub-cellular localization, as well as the identification of post-translational modifications sites in the protein.

Studies carried out by other researchers focused on the analysis of specific regions in NMNATs for other species show various functional effects depending on the intrinsic characteristics of each protein. For example, human NMNATS have a longer C-terminus of additional 24 amino acids than E. coli NaMNAT (EcNaMNAT). The bacterial enzyme was modified with an extra 17 residues, establishing that the extension of the insertion can interfere with substrate binding by decreasing the activity [36]. Other mutants for EcNaMNAT were created by making specific modifications in conserved sites close to the ATP binding motif, showing that the C-terminus is important for the stability of the protein [36]. On the other hand, studies reported by Brunetti et al. with human NMNAT-2, where four deletion mutants of different lengths were generated by eliminating from 32 to 82 residues, showed that the mutants retain the ability to synthesize NAD⁺ from NMN and ATP, demonstrating that the central domain is not essential for the functional folding of the catalytic core [37].

The existence of this type of exclusive regions in these parasite's proteins suggests that they can be involved in the regulation of the cellular mechanisms and/or control the state of operation of the enzyme, as well as in the mediation of intracellular localization functions that are

L.J. Ortiz-Joya et al.



Figure 5. 3D models and molecular dynamic simulation of LbNMNAT and $\Delta_{241-249}$ LbNMNAT mutant. (A) Structural model of LbNMNAT. The C-terminal domain (yellow), the Rossmann fold (gray), and the structural element (loop- β sheet in blue-red) connecting the α/β domain with the C-terminal domain are highlighted. The deleted region 241–249 is highlighted in red. (B) Structural overlay of the LbNMNAT (green) and $\Delta_{241-249}$ LbNMNAT (golden) models. Images generated with UCSF Chimera [33]. (C) Ligand binding site prediction in the mutated protein model. $\Delta_{241-249}$ LbNMNAT-NAD binding sites were determined by COFACTOR and COACH tools based on the structure obtained with I-TASSER (template PDB: 2H2A). Green dashed lines represent the hydrogen bonds, orange dashed line represent pi–anion interaction, and red dashed lines represent unfavorable bump interaction. BIOVIA discovery studio visualizer was used [34]. (D). Molecular dynamics simulation of LbNMNAT and $\Delta_{241-249}$ LbNMNAT protein models during 100 ns. Time evolution of backbone RMSD is shown as a function of time for the native (green) and mutant (red) structures.

essential for the correct metabolism of NAD⁺. In general, the inactivation of the 6His- $\Delta_{241-249}$ NMNAT protein from *L. braziliensis* highlights the functional importance of the identified insertion, opening an alternative to design specific inhibitors to block the parasite without affecting the human host enzymes.

Declarations

Author contribution statement

L.J.O.J designed and performed the experiments, analyzed and interpreted the data, and wrote the manuscript. L.E.C.R designed and performed the experiments, analyzed and interpreted the data. R.O. performed the experiments, analyzed and interpreted the data, and wrote the manuscript. and M.H.R.H conceived and designed the experiments, analyzed and interpreted the data, and contributed reagents and materials.

Funding statement

This work was supported by Departamento Administrativo de Ciencia, Tecnología e Innovación (COLCIENCIAS) [110165843119].

Data availability statement

Data will be made available on request.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

Supplementary content related to this article has been published online at https://doi.org/10.1016/j.heliyon.2022.e12203.

Acknowledgements

We thank the Faculty of Sciences of the Universidad Nacional de Colombia.

References

- WHO World Health Organization, OPS Organización Panamericana de la Salud, Leishmaniasis, Informe Epidemiológico de las Américas, 2018. https://iris.paho. org/bitstream/handle/10665.2/34858/LeishReport6_spa.pdf?sequence=5&is Allowed=y. (Accessed 18 May 2022).
- [2] J.D. Ramírez, C. Hernández, C.M. León, M.S. Ayala, C. Flórez, C. González, Taxonomy, diversity, temporal and geographical distribution of Cutaneous Leishmaniasis in Colombia: a retrospective study, Sci. Rep. 6 (2016), 28266.
- [3] A.M. Rezende, E.L. Folador, D. de M. Resende, J.C. Ruiz, Computational prediction of protein-protein interactions in Leishmania predicted proteomes, PLoS One 7 (2012), e51304.
- [4] R. Rojas, L. Valderrama, M. Valderrama, M.X. Varona, M. Ouellette, N.G. Saravia, Resistance to antimony and treatment failure in human Leishmania (Viannia) infection, J. Infect. Dis. 193 (2006) 1375–1383.
- [5] A. Stauch, H.-P. Duerr, J.-C. Dujardin, M. Vanaerschot, S. Sundar, M. Eichner, Treatment of visceral leishmaniasis: model-based analyses on the spread of antimony-resistant L. donovani in Bihar, India, PLoS Neglected Trop. Dis. 6 (2012), e1973.
- [6] S. Sundar, J. Chakravarty, Drug Resistance in Leishmaniasis, Antimicrobial Drug Resistance, 2017, pp. 1293–1304.
- [7] P.M. Gillespie, C.M. Beaumier, U. Strych, T. Hayward, P.J. Hotez, M.E. Bottazzi, Status of vaccine research and development of vaccines for leishmaniasis, Vaccine 34 (2016) 2992–2995.
- [8] C. Bustamante, R. Ochoa, C. Asela, C. Muskus, Repurposing of known drugs for leishmaniasis treatment using bioinformatic predictions, in vitro validations and pharmacokinetic simulations, J. Comput. Aided Mol. Des. 33 (2019) 845–854.
- [9] I.A. Rodionova, H.J. Zuccola, L. Sorci, A.E. Aleshin, M.D. Kazanov, C.-T. Ma, E. Sergienko, E.J. Rubin, C.P. Locher, A.L. Osterman, Mycobacterial nicotinate mononucleotide adenylyltransferase: structure, mechanism, and implications for drug discovery, J. Biol. Chem. 290 (2015) 7693–7706.
- [10] I. D'Angelo, N. Raffaelli, V. Dabusti, T. Lorenzi, G. Magni, M. Rizzi, Structure of nicotinamide mononucleotide adenylyltransferase: a key enzyme in NAD biosynthesis, Structure 8 (2000) 993–1004.
- [11] L. Ortiz-Joya, L.E. Contreras-Rodríguez, M.H. Ramírez-Hernández, Protein-protein interactions of the nicotinamide/nicotinate mononucleotide adenylyltransferase of Leishmania braziliensis, Mem. Inst. Oswaldo Cruz 114 (2019), e180506.
- [12] L.E. Contreras Rodríguez, M. Ziegler, M.H. Ramírez Hernández, Kinetic and oligomeric study of nicotinate/nicotinamide mononucleotide adenylyltransferase, Heliyon 6 (2020), e03733.
- [13] R. Ochoa, S.J. Watowich, A. Flórez, C.V. Mesa, S.M. Robledo, C. Muskus, Drug search for leishmaniasis: a virtual screening approach by grid computing, J. Comput. Aided Mol. Des. 30 (2016) 541–552.
- [14] R. Ochoa, A. Ortega-Pajares, F.A. Castello, F. Serral, D.F. Do Porto, J.A. Villa-Pulgarin, R.E. Varela-M, C. Muskus, Identification of potential kinase inhibitors within the PI3K/AKT pathway of Leishmania species, Biomolecules 11 (2021) 1037.
- [15] E. Gasteiger, C. Hoogland, A. Gattiker, S. 'everine Duvaud, M.R. Wilkins, R.D. Appel, A. Bairoch, Protein Identification and Analysis Tools on the ExPASy Server, the Proteomics Protocols Handbook, 2005, pp. 571–607.
- [16] N. Blom, S. Gammeltoft, S. Brunak, Sequence and structure-based prediction of eukaryotic protein phosphorylation sites, J. Mol. Biol. 294 (1999) 1351–1362.
- [17] QIAGEN CLC Genomics Workbench, Bioinformatics Software and Services QIAGEN Digital Insights, 2019 (accessed May 18, 2022), https://digitalinsights.qiagen

.com/products-overview/discovery-insights-portfolio/analysis-and-visualization /qiagen-clc-genomics-workbench/.

- [18] L.E. Contreras, R. Neme, M.H. Ramírez, Identification and functional evaluation of Leishmania braziliensis nicotinamide mononucleotide adenylyltransferase, Protein Expr. Purif. 115 (2015) 26–33.
- [19] M.M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, Anal. Biochem. 72 (1976) 248–254.
- [20] E. Balducci, M. Emanuelli, N. Raffaelli, S. Ruggieri, A. Amici, G. Magni, G. Orsomando, V. Polzonetti, P. Natalini, Assay methods for nicotinamide mononucleotide adenylyltransferase of wide applicability, Anal. Biochem. 228 (1995) 64–68.
- [21] W. Zheng, C. Zhang, Y. Li, R. Pearce, E.W. Bell, Y. Zhang, Folding non-homologous proteins by coupling deep-learning contact maps with I-TASSER assembly simulations, Cell Rep. Methods 1 (2021).
- [22] C. Zhang, P.L. Freddolino, Y. Zhang, COFACTOR: improved protein function prediction by combining structure, sequence and protein–protein interaction information, Nucleic Acids Res. 45 (2017) W291. –W299.
- [23] J. Yang, Y. Zhang, I-TASSER server: new development for protein structure and function predictions, Nucleic Acids Res. 43 (2015) W174–W181.
- [24] M. Wiederstein, M.J. Sippl, ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins, Nucleic Acids Res. 35 (2007) W407–W410.
- [25] B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation, J. Chem. Theor. Comput. 4 (2008) 435–447.
- [26] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J.L. Klepeis, R.O. Dror, D.E. Shaw, Improved side-chain torsion potentials for the Amber ff99SB protein force field, Proteins 78 (2010) 1950–1958.
- [27] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, J. Chem. Phys. 126 (2007), 014101.
- [28] M. Parrinello, A. Rahman, Crystal structure and pair potentials: a moleculardynamics study, Phys. Rev. Lett. 45 (1980) 1196–1199.
- [29] S. Stenberg, B. Stenqvist, An exact Ewald summation method in theory and practice, J. Phys. Chem. 124 (2020) 3943–3946.
- [30] D.M. Sánchez-Lancheros, L.F. Ospina-Giraldo, M.H. Ramírez-Hernández, Nicotinamide mononucleotide adenylyltransferase of Trypanosoma cruzi (TcNMNAT): a cytosol protein target for serine kinases, Mem. Inst. Oswaldo Cruz 111 (2016) 670–675.
- [31] R.G. Zhai, M. Rizzi, S. Garavaglia, Nicotinamide/nicotinic acid mononucleotide adenylyltransferase, new insights into an ancient enzyme, Cell. Mol. Life Sci. 66 (2009) 2805–2818.
- [32] J. Yang, A. Roy, Y. Zhang, Protein–ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment, Bioinformatics 29 (2013) 2588–2595.
- [33] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera-a visualization system for exploratory research and analysis, J. Comput. Chem. 25 (2004) 1605–1612.
- [34] D. Systèmes, Free Download: BIOVIA Discovery Studio Visualizer, Dassault Systèmes, 2020. https://discover.3ds.com/discovery-studio-visualizer-download. (Accessed 28 July 2022).
- [35] C. Lau, C. Dölle, T.I. Gossmann, L. Agledal, M. Niere, M. Ziegler, Isoform-specific targeting and interaction domains in human nicotinamide mononucleotide adenylyltransferases, J. Biol. Chem. 285 (2010) 18868–18876.
- [36] M. Stancek, R. Schnell, M. Rydén-Aulin, Analysis of *Escherichia coli* nicotinate mononucleotide adenylyltransferase mutants in vivo and in vitro, BMC Biochem. 6 (2005) 16.
- [37] L. Brunetti, M. Di Stefano, S. Ruggieri, F. Cimadamore, G. Magni, Homology modeling and deletion mutants of human nicotinamide mononucleotide adenylyltransferase isozyme 2: new insights on structure and function relationship, Protein Sci. 19 (2010) 2440–2450.