Journal of Molecular Graphics and Modelling 27 (2009) 701–705

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Islet neogenesis associated protein (ingap): Structural and dynamical properties of its active pentadecapeptide

Andrés N. McCarthy [a,b,*], Inés G. Mogilner [a], J. Raúl Grigera [a,b], M. Inés Borelli [c], Hector Del Zotto [c], Juan José Gagliardino [c]

[a] Instituto de Física de Líquidos y Sistemas Biológicos, (CONICET La Plata-UNLP-CIC) 59N° 789, c.c. 565, B1900BTE La Plata, Buenos Aires, Argentina
[b] Departamento de Ciencias Biológicas, Facultad de Ciencias Exactas, Universidad Nacional de La Plata, 47 y 115, (1900) La Plata, Buenos Aires, Argentina
[c] CENEXA, Centro de Endocrinología Experimental y Aplicada, (UNLP-CONICET La Plata), Facultad de Ciencias Médicas, UNLP. Calle 60 y 120, 1900 La Plata, Buenos Aires, Argentina

ARTICLE INFO

Article history:
Received 23 April 2008
Received in revised form 28 October 2008
Accepted 1 November 2008
Available online 8 November 2008

Keywords:
Molecular dynamics
Conformation
Neogenesis associated protein
Diabetes

ABSTRACT

We have studied the structural and dynamical properties of the biologically active pentadecapeptide of the islet neogenesis associated protein (INGAP-PP) and of two other pentadecapeptides with the same aminoacid composition but randomly scrambled primary sequences, using molecular dynamic simulations. Our data demonstrates that whilst the peptides with scrambled sequences show no definite prevalent structure in solution, INGAP-PP maintains a notably stable tertiary fold, namely, a conformer with a central β-sheet and closed C-terminal. Such structure resembles the one corresponding to the aminoacid sequence of human pancreatitis associated protein-1 (PAP-1), which presents 85% sequence homology with INGAP. These results could reasonably explain why the two scrambled sequences tested showed no biological activity, while INGAP-PP significantly increases β-cells function and mass both *in vitro* and *in vivo* conditions. The capability of INGAP-PP to temporarily adopt other closely related conformations offers also a plausible explanation for the 50 fold experimental difference in potency between the active pentadecapepide and the whole protein. They also suggest that the C-terminal region of INGAP-PP may plausibly be the locus for its interaction with the cell receptor. Consequently, the knowledge gathered through our data can help to obtain more potent INGAP-PP analogs, suitable for the prevention and treatment of diabetes.

1. Introduction

Several genes, hormones and growth factors participate in the control of the development and growth of pancreatic β-cells [1]. Several evidences demonstrate that islet neogenesis associated protein (INGAP) is one of these factors.

INGAP was a compound originally identified as part of a protein complex (ilotropin), isolated from pancreata of normal adult hamsters previously wrapped in cellophane [2], and thereafter cloned and sequenced [3]. Lately however, the presence of INGAP was found at the embryonic mouse pancreas commitment period, thus providing evidence of its early normal presence and possible role in pancreas development and patterning [4].

INGAP gene encodes a 19,940 Da transcript [3], only produced in different subsectors of the pancreas (islets, duct and exocrine cells) [5]. A synthetic pentadecapeptide having the 104–118 aminoacids sequence of INGAP (INGAP-PP), reproduces the stimulatory effect of the intact molecule upon thymidine incorporation into pancreatic duct cells and duct-cell lines [3]. Using a model of insulin resistance induced by dietary manipulation, we have consistently shown a simultaneous and significant increase in β-cell mass, glucose-induced insulin secretion [6,7] and the appearance of a possible early islet precursor cell co-expressing INGAP/Pdx-1 [8].

On the other hand, it has been shown that injection of INGAP-PP to either normal or streptozotocin-induced diabetic mice is accompanied by an increase in β-cell mass and signs of islets neogenesis [9]. Further, we have also found that neonatal and adult rat islets cultured with INGAP-PP release more insulin in response to glucose and other secretagogues [10] and increase the expression of genes involved in the process of insulin secretion [11]. Altogether these results suggest that INGAP plays a role in the control of islets development and growth not only in the fetal but also in the adult period as well.

^{125}I-labelled analog of INGAP-PP, injected into normal hamsters, specifically binds to pancreas, liver and small bowel tissues

* Corresponding author. Tel.: +54 221 4233283/4254904; fax: +54 221 4257317.
E-mail address: amccarthy@iflysib.unlp.edu.ar (A.N. McCarthy).

1093-3263/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2008.11.001

[12]. However, neither the action mechanism nor the possible cell receptor structures for either the INGAP or INGAP-PP molecules have been so far identified.

In order to gain some insight into the molecular mechanism involved in the above mentioned effects, we have performed structural and dynamic studies of the synthetic INGAP-PP in solution, using molecular dynamics simulation. We have simultaneously studied several other randomly scrambled sequences of INGAP-PP to use as negative controls. The results obtained were compared with the structure of pancreatitis-associated protein 1 currently available in the literature (PDB entries 1UV0 and 2GO0), which has an 85% sequence homology with INGAP-PP.

Our data show that whilst the scrambled sequences (biologically inactive), randomly present a virtually limitless conformational space (not converging to any given structure), the original sequence (biologically active), mostly alternates between three very distinct conformations.

2. Computational method

We carried out molecular dynamic (MD) simulations using the GROMACS 3.2.1 package [13–16], in which the equations of motion are solved using a leap-frog integration step. We used all-atoms force field [17–22] for the minimization process, as well as for the entire MD simulation stages and kept all protein bond lengths and water molecules constrained using the LINCS algorithm [23] and the SETTLE algorithm [24], respectively. We calculated the electrostatic forces using the Reaction Field method while the Lennard–Jones interactions were calculated within a cut-off radius of 1.4 nm.

We ran all the simulations using a Xeon based, dual-processor cluster and the GNU/Linux, while all plots and graphics were performed using MS Windows or GNU/Linux and the reference Visual Molecular Dynamics package, Swiss PDB Viewer or XGrace software.

As starting configurations of each pentadecapeptide sequence (INGAP-PP and scrambled peptides), we have used structures produced by the Swiss PDB Viewer software, with phi–psi (φ, ψ) angles of (-180°, 180°). INGAP-PP average structures were compared to experimentally resolved structures of the pancreatitis-associated protein 1 currently available in literature (PDB entries 1UV0 and 2GO0). Topologies were generated using the PDB2GMX tool, with standard pH 7.0 aminoacid protonation states. The SPC/E [25] water model was used for solvating all systems. All starting systems consisted of a cubic simulation box of $X = 7.19617$ nm, $Y = 7.19617$ nm and $Z = 7.19617$ nm, with a total volume of 372.653 nm^3, containing one pentadecapeptide molecule and 11,663 water molecules. All systems were energy minimized using initially the steepest descent method, converging to machine precision, and then applying the conjugated gradient method, converging in less than 20 cycles. Thereafter, the solvent was allowed to relax running a short 10 ps MD simulation with position restraints applied on all the pentadecapeptide heavy atoms. At the beginning all systems were weakly coupled to a thermal and hydrostatic bath, in order to work within the isothermal–isobaric ensemble [26] at $T = 300$ K and $P = 1$ bar. Subsequently, the whole system was allowed to relax (restraints were removed from all atoms) during 200 ps, coupled to the same weak thermal and hydrostatic bath. At this point and in order to produce comparable equilibrated starting points for every system, all systems were allowed to stabilize during a simulation period of 10 ns. Once this equilibrium was achieved, we performed a 20 ns long MD simulation for all systems. The analysis of the three pentadecapeptide sequences was performed using these 20 ns long MD simulations. For graphical clearness, the data presented corresponds to INGAP-PP and the scrambled sequence which showed the smallest structural divergence throughout the whole simulated time. In order to confirm the behavior observed during the first 20 ns, the simulations on these two systems were extended for a further 30 ns, reaching a total simulated time of 50 ns. The data obtained from this extension is presented in the Supplementary Material.

All plots and matrices were calculated using the corresponding GROMACS package programs.

3. Results and discussion

3.1. Root mean square deviation (RMSD) analysis

The initial structure from each MD data collection run was used as a reference for the calculation of its corresponding run alpha carbon RMSD. Prior to this, the peptide structures from every recorded frame in all trajectories were progressively least square fitted, thus removing the effects of rotation and translation. To effectively compare the relative structural and dynamical properties of the INGAP-PP and of the scrambled peptide systems, all RMSD were plotted using the same scales, both in the color coded scale and in nanometers (nm). In order to enhance graphical clearness, 200 ps running averages were performed on the RMSD curves.

Fig. 1 shows the RMSD for both the INGAP-PP and the scrambled peptide systems, from 0 ns to 20 ns. The analysis of this 20 ns period of MD simulation time shows a discrete fluctuation of the RMSD value for the INGAP-PP system. In fact, whilst during the whole 20 ns period this system appears to have the ability to explore structurally related conformations, it returns frequently to an apparently conserved tertiary fold. On the contrary, the scrambled peptide systems showed a completely different behavior: RMSD values neither oscillate around any given value, nor return to any given conformation, thus, showing a behavior compatible with a random structural exploration adopting no prevalent conformer. Although such observations are expected for such short peptides in solution, they contrast with the behavior of the original INGAP-PP, stressing consequently the importance of the structural stability shown by the latter.

Based on these data, we can conclude that under identical experimental conditions, whilst the INGAP-PP system reaches a point of reasonable stability, around which it fluctuates, the scrambled systems show a much more unstable behavior. The validity of these conclusions has been corroborated by extending

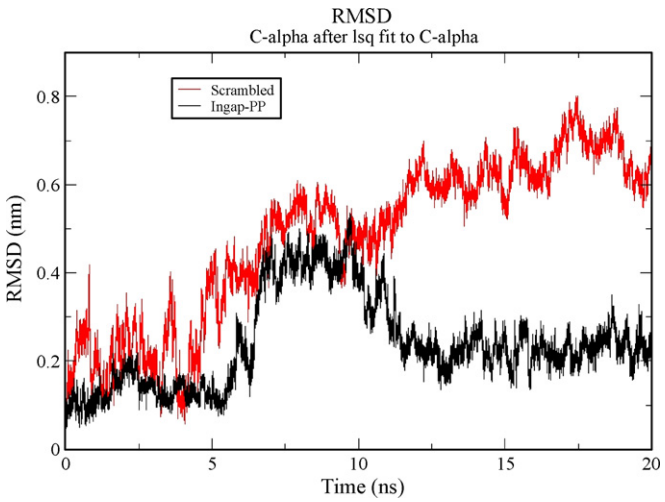

Fig. 1. Data represents the absolute root mean square deviation (RMSD) for both the INGAP-PP system (black) and the scrambled systems (red) from 0 to 20 ns.

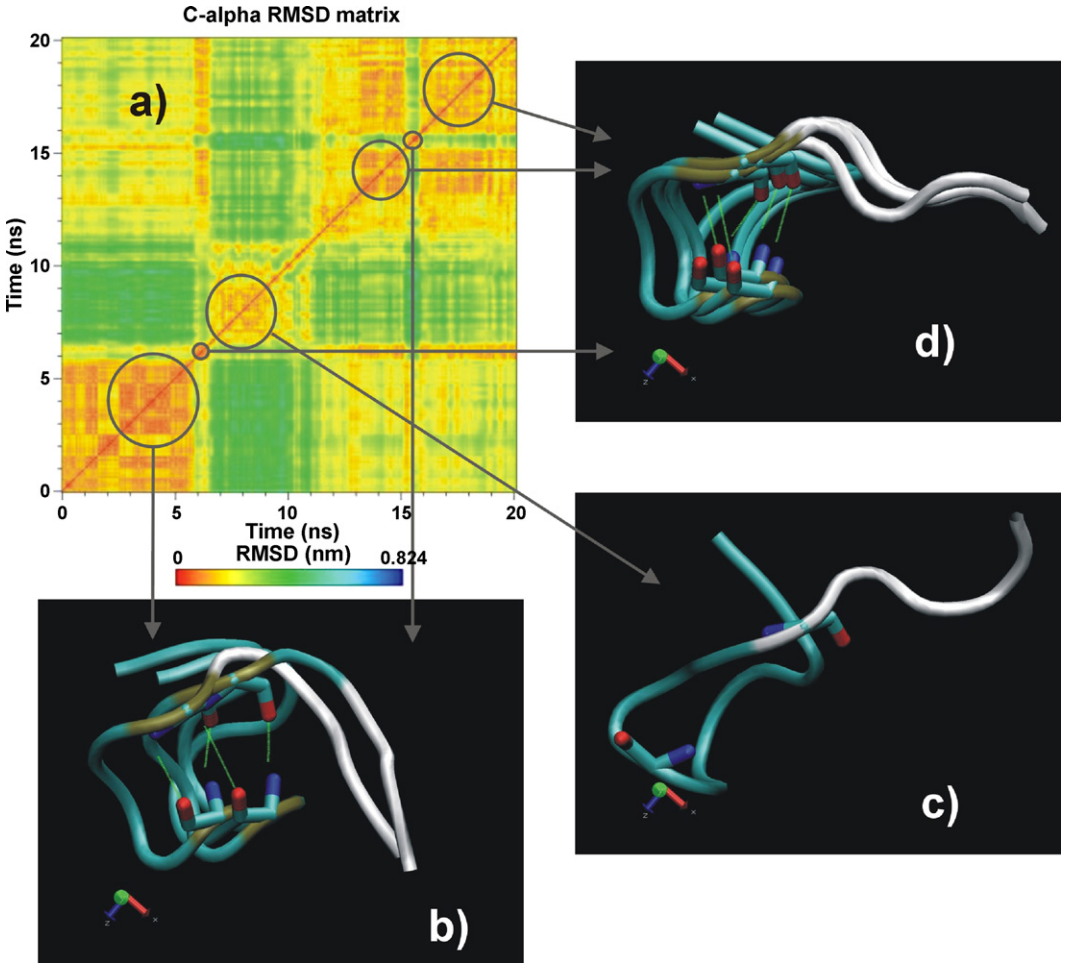

Fig. 2. (a) Shows the root mean square deviation (RMSD) matrix for the whole of the INGAP-PP system simulation time (from 0 to 20 ns). The RMSD value was calculated pair wise, covering every recorded frame combination. The results are presented in a color code format that can be readily translated into nanometer (nm). (b) Least square fitted ribbon representation of the INGAP-PP structure, averaging over all frames collected for the simulated periods of 2.5 to 4.7 ns and 15.3 to 15.8 ns. (c) Ribbon representation of the INGAP-PP structure, averaging over all frames collected for the simulated period 7 to 9 ns. (d) Least square fitted ribbon representation of the INGAP-PP structures, averaging over all frames collected for the simulated periods of 5.8 to 6.5 ns, 13 to 15 and 16.5 to 18.5 ns. Secondary structure color codes for (b–d): cyan → turn, white → coil and golden → β-sheet. The dotted green lines show the existence of backbone hydrogen bonds.

the simulation time (and the corresponding RMSD analysis) of these two systems for an extra 30 ns (see Supplementary Material).

Fig. 2(a) and Fig. 3 offer a more detailed approach to this issue. In order to keep the color coded limits representative between simulations, in Fig. 2(a) the reference color code was scaled to match that of Fig. 3. This procedure allows for a direct graphical comparison of the results from the INGAP-PP and the scrambled sequence runs.

Fig. 2(a) shows three stable dominant (orange-yellow) regions, connected by intermediate regions of slightly higher (yellow-green) fluctuation values. This graphical data suggests that most fluctuations in the matrix keep within low RMSD values, being thus compatible with a homogeneous tertiary fold. Further, it supports the assumption gathered from Fig. 1 that the INGAP-PP MD simulation has reached equilibrium, and explores structures which are mostly closely related.

Likewise, from a general graphical analysis of Fig. 3, we may conclude that the scrambled sequence structures do not tend to repeat themselves. In fact, even during the short periods of apparent stability (represented by the yellow-green regions), most structures differ markedly from each other. The comparison reaches deviation values which almost double any of the RMSD values observed in Fig. 2 (blue regions).

Looking at the data collected from the simulation for the INGAP-PP sequence, the three orange yellow regions may be readily separated into two distinct groups. The first and statistically more predominant group shows a conserved general fold comprised by the occurrence of a central β-sheet, flanked by a turn on the N-terminal side and a random coil on the C- terminal side (Fig. 2(b) and (d)). On the contrary, the second group is characterized by the loss of the central β-sheet, which appears to lead to a disruption of the previously adopted fold (Fig. 2(c)). Likewise, the first group may be divided into two subgroups, which differ only in the open (Fig. 2(d)) or closed (Fig. 2(b)) position of the more mobile C-terminal random coil. It is interesting to observe that during the simulated time, INGAP-PP shows the ability to move reversibly between these three structures. The structures for these three distinct INGAP-PP conformers are available in the Supplementary Material.

In order to explore the possible similarities between simulation results and those experimentally determined for structures with homologous sequences, we performed a BLAST sequence alignment for the INGAP-PP 15 aminoacids sequence against all proteins available in the sequence database, as well as in the protein data bank (Table 1). The complete list for the top first five BLAST matches is provided in the Supplementary Material.

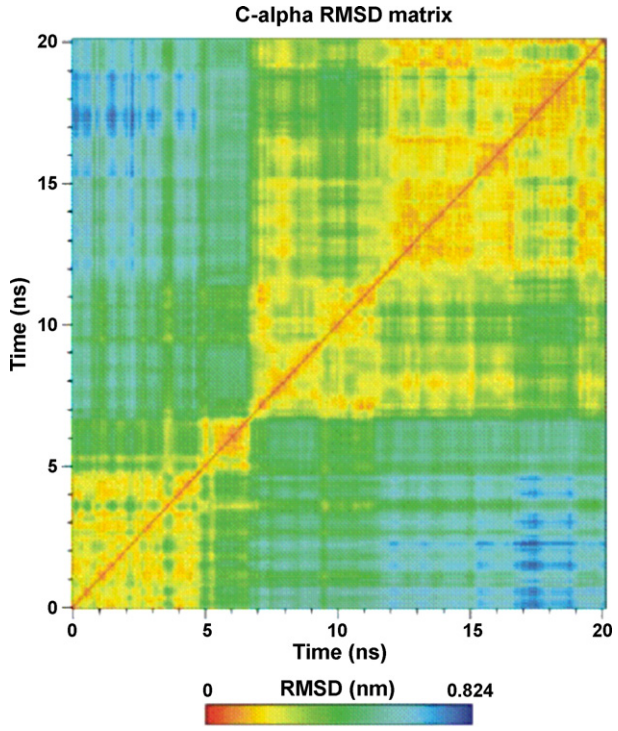

Fig. 3. Shows the root mean square deviation (RMSD) matrix for the whole time-period of the scrambled systems simulation (from 0 ns to 20 ns). As in Fig. 2, the RMSD value was calculated pair wise, covering every recorded frame combination. The results are presented in a color code format that can be readily translated into nanometer (nm).

Table 1 shows the results of the alignment procedure. It may be observed that human pancreatitis-associated protein1 (PAP1) and INGAP-PP present the best fitted alignment for proteins with experimentally resolved tertiary structure. Furthermore, it is the best overall alignment outside the INGAP protein family. Bearing in mind that the alignment was performed over the complete database, it is remarkable to find such a notorious sequential

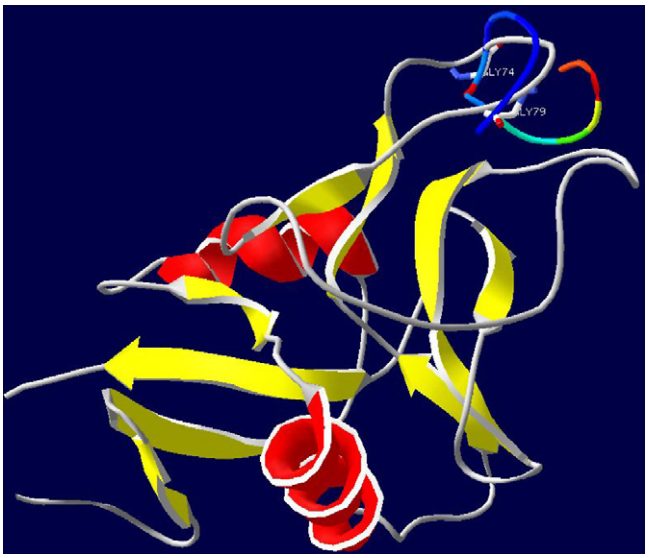

Fig. 4. Shows the best least square fit between all the average structures presented in Fig. 2 and the available experimental structure for PAP-1 (PDB code 1GV0), which corresponds to the average structures represented in Fig. 2(b).

Table 1
Best result for BLAST sequence alignment of INGAP-PP against sequences with experimentally resolved tertiary structure.

pdb|2GO0|A Chain A, Nmr solution structure of human pancreatitis-associated protein		
Length = 137		
Query 1	IGLHDPSHGTLPNG	14
	IGLHDP + GT PNG	
Subject 66	IGLHDPTQGTEPNG	79
Score = 34.6 bits (74), expect = 0.78, method: composition-based stats		
Identities = 11/14 (78%), positives = 12/14 (85%), gaps = 0/14 (0%)		

similarity between the peptide under study and a neogenetic marker, such as the PAPI. Further still, one of the stable conformers also presents a high degree of structural similarity. Although notable, these last results are very much in line with recent publications which strongly correlate these two endocrine (diabetes) and exocrine (pancreatitis) markers [27].

We subsequently performed a least squares fit for every average structure presented in Fig. 2 against the referred experimental structure for PAP1, obtaining the best fit for the average structures shown in Fig. 2(b). A graphical result of such a fit can be seen in Fig. 4, which shows a very good structural resemblance between the simulated and the experimental structure. It is interesting to note that within this structural alignment, whilst the best fit corresponds to the C-terminal aminoacids, those corresponding to the N-terminal region seem to occupy the position of the rest of the missing protein, thus creating the computationally observable stable fold.

The C-terminal portion of INGAP-PP is not only the most structurally stable component, but it is also the portion of the polypeptide structure which bears the highest resemblance with the experimental structure. Hence, we can assume that this aminoacid sequence might be a suitable candidate to interact with the INGAP-PP receptor. The study of the biological activity of molecules with sequential blockage or aminoacid replacement of this locus could provide a clear evidence to support this assumption.

4. Conclusions

Our data demonstrates that only the active INGAP-PP peptide exhibits some stable and predictable structure in solution. It also suggests that the C-terminal region of INGAP-PP may plausibly be the locus for its interaction with the cell receptor. Consequently, the current results will help to understand the INGAP-PP interaction with its specific receptor/s and its action mechanism. On time, this knowledge could also help to develop more potent INGAP-PP analogs, maintaining the structure of the central β-sheet and the closed C-terminal conformer, suitable for its use in the prevention and treatment of diabetes.

Acknowledgements

This work was partially supported by the Consejo Nacional de Investigaciones Científicas y Técnicas of Argentina (CONICET), the Universided de La Plata (UNLP), the Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CICPBA) and the Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT). JRG, MIB, HdZ and JJG are members of the Carrera del Investigador of CONICET, ANMcC is supported by UNLP, and is a member of the Carrera del Investigador of CICPBA.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmgm.2008.11.001.

References

[1] J.H. Nielsen, E.D. Galsgaard, A. Møldrup, B.N. Friedrichsen, N. Billestrup, J.A. Hansen, Y.C. Lee, C. Carlsson, Regulation of β-cell mass by hormones and growth factors, Diabetes 50 (Suppl. 1) (2001) S25–S29.

[2] G.L. Pittenger, A.I. Vinik, L. Rosenberg, The partial isolation and characterization of ilotropin, a novel islet-specific growth factor, Adv. Exp. Med. Biol. 321 (1992) 123–130.

[3] R. Rafaeloff, G.L. Pittenger, S.W. Barlow, X.F. Qin, B. Yan, L. Rosenberg, W.P. Duguid, A.I. Vinik, Cloning and sequencing of the pancreatic islet neogenesis associated protein (INGAP) gene and its expression in islet neogenesis in hamsters, J. Clin. Invest. 9 (1997) 09–2100.

[4] N.S. Hamblet, W. Shi, A.I. Vinik, D.A. Taylor-Fishwick, The Reg family member INGAP is a marker of endocrine patterning in the embryonic pancreas, Pancreas 36 (1–9) (2008).

[5] L.E. Flores, M.E. García, M.I. Borelli, H. Del Zotto, M.E. Alzugaray, B. Maiztegui, J.J. Gagliardino, Expression of islet neogenesis-associated protein (INGAP) in islets of normal hamsters., J. Endocrinol. 177 (2003) 243–248.

[6] H. Del Zotto, L. Massa, C.L. Gómez Dumm, J.J. Gagliardino, Changes induced by sucrose administration upon the morphology and function of pancreatic islets in the normal hamster., Diabetes Metab. Res. Rev. 15 (1999) 106–112.

[7] H. Del Zotto, L. Massa, R. Rafaeloff, G.L. Pittenger, A. Vinik, G. Gold, A. Reifel-Miller, J.J. Gagliardino, Possible relationship between changes in islet neogenesis and islet neogenesis-associated protein-positive cell mass induced by sucrose administration to normal hamsters., J. Endocrinol. 165 (2000) 725–733.

[8] J.J. Gagliardino, H. Del Zotto, L. Massa, L.E. Flores, M.I. Borelli, Pancreatic duodenal homeobox-1 (Pdx-1) and islet neogenesis associated protein (INGAP): a possible combined marker of activatable pancreatic cell precursors., J. Endocrinol. 177 (2003) 249–259.

[9] L. Rosenberg, M. Lipsett, J.W. Yoon, M. Prentki, R. Wang, H.S. Jun, G.L. Pittenger, D. Taylor-Fishwick, A.I. Vinik, A pentadecapeptide fragment of islet neogenesis-associated protein increases beta-cell mass and reverses diabetes in C57BL/6J mice., Ann. Surg. 240 (2004) 875–884.

[10] M.I. Borelli, L.F. Stoppiglia, L.F. Rezende, L.E. Flores, H. Del Zotto, A.C. Boschero, Gagliardino JJ:, INGAP-related pentadecapeptide: its modulatory effect upon insulin secretion, Regul. Pept. 131 (2005) 97–102.

[11] H. Barbosa, S. Bordin, L.F. Stoppiglia, K.E. Silva, H. Del Zotto, J.J. Gagliardino, Boschero AC:, Islet neogenesis associated protein (INGAP) modulates gene expression in cultured neonatal rat islets, Regul .Pept. 136 (2006) 78–84.

[12] M.I. Borelli, H. Del Zotto, L.E. Flores, M.E. García, A.C. Boschero, Gagliardino JJ:, Transcription, expression and tissue binding *in vivo* of INGAP and INGAP-related peptide in normal hamsters, Regul Pept. 140 (2007) 192–197.

[13] H.J.C. Berendsen, D. van der Spoel, R. van Drunen, GROMACS: a message-passing parallel molecular dynamics implementation, Comput. Phys. Commun. 91 (1995) 43–56.

[14] E. Lindahl, B. Hess, D. van der Spoel, Gromacs 3.0: a package for molecular simulation and trajectory analysis, J. Mol. Mod. 7 (2001) 306–317.

[15] W.F. van Gunsteren, H.J.C. Berendsen, Computer simulation of molecular dynamics: methodology, applications, and perspectives in chemistry, Angew. Chem. Int. Ed. Engl. 29 (1990) 992–1023.

[16] D. van der Spoel, E. Lindahl, B. Hess, A.R. van Buuren, E. Apol, P.J. Meulenhoff, D.P. Tieleman, A.L.T.M. Sijbers, K.A. Feenstra, R. van Drunen, H.J.C. Berendsen, Gromacs User Manual version 3.2, www.gromacs.org, 2004.

[17] W.F. van Gunsteren, H.J.C. Berendsen, Gromos-87 manual, Biomos BV Nijenborgh 4, Groningen, The Netherlands, 1987.

[18] A.R. van Buuren, S.J. Marrink, H.J.C. Berendsen, A molecular dynamics study of the decane/water interface, J. Phys. Chem. 97 (1993) 9206–9212.

[19] A.E. Mark, S.P. van Helden, P.E. Smith, L.H.M. Janssen, W.F. van Gunsteren, Convergence properties of free energy calculations: cyclodextrin complexes as a case study, J. Am. Chem. Soc. 116 (1994) 6293–6302.

[20] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 926–935.

[21] A.R. van Buuren, H.J.C. Berendsen, Molecular dynamics simulation of the stability of a 22 residue alpha-helix in water and 30% trifluoroethanol., Biopolymers 33 (1993) 1159–1166.

[22] H. Liu, F. Müller-Plathe, W.F. van Gunsteren, A force field for liquid dimethyl sulfoxide and liquid proporties of liquid dimethyl sulfoxide calculated using molecular dynamics simulation, J. Am. Chem. Soc. 117 (1995) 4363–4366.

[23] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, LINCS A linear constraint solver for molecular simulations, J. Comp. Chem. 18 (1997) 1463–1472.

[24] S. Miyamoto, P.A. Kollman, SETTLE An analytical version of the SHAKE and RATTLE algorithms for rigid water models, J. Comput. Chem. 13 (1992) 952–962.

[25] H.J.C. Berendsen, J.R. Grigera, T.P. Straatsma, The missing term in effective pair potentials, J. Chem. Phys. 91 (1987) 6269–6271.

[26] H.J.C. Berendsen, J.P.M. Postma, A. DiNola, J.R. Haak, Molecular dynamics with coupling to an external bath, J. Chem. Phys. 81 (1984) 3684–3690.

[27] R. Graf, M. Schiesser, Reding Th, P. Appenzeller, L.-K. Sun, F. Fortunato, A. Perren, D. Bimmler, Exocrine meets endocrine: pancreatic Stone Protein-Two sides of the same coin, J. Surg. Res. 133 (2006) 113–120.