Research Article

A series of Notch3 mutations in CADASIL; insights from 3D molecular modelling and evolutionary analyses

Dimitrios Vlachakis[1#], Spyridon Champeris Tsaniras[2#], Katerina Ioannidou[3], Louis Papageorgiou[1], Marc Baumann[4] and Sophia Kossida[1]

[#] These authors contributed equally to this study
[1] Bioinformatics & Medical Informatics Team, Biomedical Research Foundation, Academy of Athens, Athens, Greece
[2] Department of Physiology, Medical School, University of Patras, Rio, 26504 Patras, Greece
[3] School of Electrical and Computer Engineering, National Technical University of Athens, Greece
[4] Protein Chemistry/Proteomics Unit, Biomedicum Helsinki, Institute of Biomedicine, University of Helsinki, P.O. Box 63, Finland

Received on October 5, 2014; Accepted on October 20, 2014; Published on October 31, 2014

Correspondence should be addressed to Dimitrios Vlachakis (dvlachakis@bioacademy.gr) and Sophia Kossida (skossida@bioacademy.gr)

Abstract

CADASIL disease belongs to the group of rare diseases. It is well established that the Notch3 protein is primarily responsible for the development of CADASIL syndrome. Herein, we attempt to shed light to the actual molecular mechanism underlying CADASIL via insights that we have from preliminary *in silico* and proteomics studies on the Notch3 protein. At the moment, we are aware of a series of Notch3 point mutations that promote CADASIL. In this direction, we investigate the nature, extent, physicochemical and structural significance of the mutant species in an effort to identify the underlying mechanism of Notch3 role and implications in cell signal transduction. Overall, our *in silico* study has revealed a rather complex molecular mechanism of Notch3 on the structural level; depending of the nature and position of each mutation, a consensus significant loss of beta-sheet structure is observed throughout all *in silico* modeled mutant/wild type biological systems.

Introduction

CADASIL stands for Cerebral Autosomal Dominant Arteriopathy Subcortical Infarcts Leukoencephalopathy. In other words, cerebral relates to the brain, autosomal dominant to the inheritance pattern, arteriopathy to arterial disease (in this case within the brain), subcortical to the area of the brain involved in higher functioning, infarcts to tissue necrosis due to a lack of oxygen and leukoencephalopathy to myelin damage in the brain (Joutel *et al.* 1996). CADASIL is also known as hereditary multi-infarct dementia, familial disorder with subcortical ischemic strokes and familial Binswanger's disease.

This disease is the most prevalent type of hereditary cerebral angiopathy (Sourander & Walinder 1977) or stroke disorder, and mutations of the NOTCH3 gene located on chromosome 19 are regarded as its main cause. It belongs to a family of disorders called leukodystrophies (Chabriat *et al.* 1995, Joutel *et al.* 1996). More than 30 years have passed since the first time the condition was described in a Swedish family. However, the acronym CADASIL did not come out until the early 1990s (Sourander & Walinder 1977, Tournier-Lasserve *et al.* 1993). Most patients with CADASIL have a family member with the disorder. However, there have been cases reported where a patient has no first-degree relative with CADASIL symptoms (Joutel *et al.* 2000). Nevertheless, as genetic testing for this disease was developed after 2000, many cases were misdiagnosed as different neurodegenerative disorders (Razvi *et al.* 2005).

This inherited condition mainly affects the small arteries in the white matter of the brain, causing thickening of their walls that, consequently, blocks the blood flow. The poorer blood supply can cause areas of tissue death; for example, multiple lacunar and subcortical white matter infarctions. The vascular smooth muscle (VSM) cells surrounding the arteries become abnormal and gradually die. Disproportionate cortical hypometabolism has also been reported (Tatsch *et al.* 2003).

The most common first sign of CADASIL is migraines, often followed by visual sensations or auras, usually at the age of 20-30. However, the most usual symptom, which has also been reported as a first sign in some cases, is the ischemic episodes, which include Transient Ischemic Attacks (TIAs) and strokes. Strokes appear for the first time between the ages of 30 and 50, while TIAs usually occur prior to

them. Individuals with CADASIL may suffer recurrent strokes (subcortical cerebral infarctions) (Sourander & Walinder 1977) throughout their lifetime that damage the brain as time goes by. Strokes can cause paralysis like pseudobulbar palsy (Chabriat *et al.* 1995, Dalkas *et al.* 2013), loss of sensation, walking problems or slurred speech, while recurrent ischemic episodes may lead to urinary incontinence or severe disability. The strokes that appear in the subcortical area of the brain, associated with reasoning and memory, are also responsible for cognitive function problems. Cognitive deterioration progresses to difficulties with concentration, attention or loss of intellectual function (Toni-Uebari 2013). Approximately 70% of the patients over 65 years old suffer from subcortical dementia, which appears with slowing of motor function, loss of memory or apathy. Furthermore, almost 10-20% of the patients demonstrate psychiatric disorders, sometimes even as a first symptom. Such disorders include mood or perception changes (mania or depression), changes in personality, hallucinations, delusions, anxiety and panic disorders (Valenti *et al.* 2011). Rarely, CADASIL also causes recurrent epileptic seizures (5-10% of patients) (Buffon *et al.* 2006, Dichgans *et al.* 1998, Velizarova *et al.* 2011). Leukoencephalopathy is diagnosed through magnetic resonance imaging (MRI) as diffuse white matter lesions (Ueda *et al.* 2009).

The onset of cerebrovascular disease usually occurs in the mid-adult (30s-60s), while recurrent transient ischemic attacks (TIAs) or strokes in multiple vascular territories appear between 30-50 years of age. Symptoms and disease onset can vary widely (Chabriat et al. 1995).

The Notch3 Protein

In general, Notch genes encode receptors that arbitrate short-range signalling events. A typical Notch gene encodes a single-pass transmembrane receptor protein (Fiuza & Arias 2007, Sourander & Walinder 1977).

In particular, the NOTCH3 gene was identified in the early 90s and its expression was first identified in proliferating neuroepithelium (Bellavia *et al.* 2008). It provides instructions for generating the Notch3 receptor protein, found on the surface of VSM cells (Tikka *et al.* 2009). Upon Notch3 receptor activation, a signaling cascade is initiated resulting in the regulation of specific genes. In the case of VSM cells, Notch3 receptors play a pivotal role in the optimal functionality, survival and maintenance of muscle cells in the arterial network of the brain.

The NOTCH3 gene is located on the short (p) arm of chromosome 19 between positions 13.2 and 13.1 (cytogenetic location), as retrieved from NCBI database (Gene ID: 4854). More precisely, the molecular location of NOTCH3 gene is between the base pairs 15,270,443 to 15,311,791 on chromosome 19.

Methods

Bioinformatics

A Blast search, with the Notch 3 protein sequence was performed at the European Bioinformatics Institute (EBI), to find similar sequences in the Uniprot database (Consortium 2011). All sequences with an E-value lower than 7.0E-4 were selected to perform an analysis with the Multiple EM for Motif Elicitation (MEME) software (Papageorgiou *et al.* 2014, Papangelopoulos *et al.* 2014). The analysis was performed at http://meme.sdsc.edu/meme4_6_1/ using default values. Logos for the found motifs were generated using the WebLogo software (Crooks *et al.* 2004).

Medical Data

A list of point mutations of the Notch3 protein was gracefully provided by Prof. Marc Baumann, the Biomedicum Helsinki Unit. Table 1 summarizes the point mutations examined in this study.

Table 1. Point mutations examined in the current study.

A	Arg107Trp	Canadian case, previously unreported (data provided by Prof. Baumann)
B	Ser497Leu	Reported in the ESP database (NHLBI GO Exome Sequencing Project) (data provided by Prof. Baumann)
C	Glu813Lys	Reported in the ESP database (NHLBI GO Exome Sequencing Project) (data provided by Prof. Baumann)
D	Ala1020Pro	Scheid *et al.* 2008
E	Ser978Arg	Ferreira *et al.* 2007
F	His1133Glu	Canadian case (data provided by Prof. Baumann)

Homology Modelling and Energy Minimization

Homology modelling for the Notch3 protein was carried out using Modeller (Sali 1995). The structure of the Human Notch1 EGFS structure was used as template for this study (PDB entry: 2VJ3) (Cordle *et al.* 2008). Subsequent energy minimization was performed using the Gromacs-implemented, Charmm27 forcefield (Vlachakis *et al.* 2013a, b). Models were structurally evaluated using the Procheck utility (Laskowski 1996, Vlachakis *et al.* 2013c). Energy minimizations were used to remove any residual geometrical strain in each molecular system, using the Charmm forcefield as it is implemented into the Gromacs suite, version 4.5.5 (Hess *et al.* 2008, Pronk *et al.* 2013). All Gromacs-related simulations were performed through our previously developed graphical interface (Sellis *et al.* 2009, Vlachakis *et al.* 2013d). An implicit Generalized Born (GB) solvation was chosen at this stage, in an attempt to speed up the energy

minimization process (Vlachakis & Kossida 2014).

Molecular Dynamics Simulations

Molecular systems of the Notch3 protein were subjected to unrestrained Molecular Dynamics simulations (MDs) using the Gromacs suite, version 4.5.5 (Hess *et al.* 2008, Pronk *et al.* 2013, Vlachakis *et al.* 2014). MDs took place in a SPC water-solvated, periodic environment. Water molecules were added using the truncated octahedron box extending 7Å from each atom. Molecular systems were neutralized with counter-ions as required. For the purposes of this study all MDs were performed using the NVT ensemble in a canonical environment, at 300 K, 1 atm and a step size equal to 2 femtoseconds for a total 500 nanoseconds simulation time (Balatsos *et al.* 2012, Vlachakis *et al.* 2012b, 2013f, g). An NVT ensemble requires that the number of atoms, volume and temperature remain constant throughout the simulation (Vlachakis *et al.* 2012a, 2013e).

Results & Discussion

Molecular Mechanism

More than 190 mutations in the NOTCH3 gene have been found to cause CADASIL (Tournier-Lasserve *et al.* 1993). The gene encodes a large transmembrane (Notch3) receptor protein (Joutel *et al.* 1996), which has a crucial role within the vascular smooth muscle signaling pathways. It is comprised of 33 exons (Monet-Leprêtre 2009). At the molecular level, all pathogenic mutations have been found to result in an odd number of cysteine residues (which stay unpaired, due to addition or deletion of a cysteine molecule) in one of the 34 epidermal growth factor (EGF)-like repeats that are present in the extracellular domain (N3ECD). This results in misfolding of the Notch3 protein. NOTCH3 is expressed in VSM cells and this misfolding leads to a gradual accumulation of N3ECD on their surface, resulting in their degeneration. Apart for N3ECD accumulation, there is also granular osmiophilic material (GOM) deposited in the walls of the vessels. These two features can only be found in CADASIL (Ishiko *et al.* 2006, Lewandowska *et al.* 2011, Valenti *et al.* 2011).

Various neurodegenerative diseases have presented the key pathological feature of aggregation and accumulation of misfolded proteins. It was recently reported that Notch3 multimerization is enhanced by the presence of CADASIL-associated mutations. In the case that GOM deposits contain N3ECD, the mutated Nothc3 protein oligomers that are formed may be the lodgements of GOM. Degradation does not affect mutant Notch3 aggregates and their formation may cause cellular dysfunction and/or cell death. However, VSM cell degeneration and loss may not be the only cause of blood vessel wall deterioration in CADASIL patients. Other factors contributing to this process may include hyalinization of the vessels, fibrosis and an abnormal increase in the intercellular space at the vessel wall (Ishiko *et al.* 2006, Lewandowska *et al.* 2011, Valenti *et al.* 2011).

However, recently some non-Cys mutations have been reported. Those mutations do not follow the characteristic pathology and pattern of the disease. Furthermore, it has been identified that some of the non-Cys mutations cause GOM whereas others do not (Bersano *et al.* 2012).

AAR Pattern

Amino Acid Repeats (AARs) can be frequently found within protein sequences and play various roles in protein function and evolution. Therefore, the identification of AAR patterns usually serves as the first step for further investigation. Various algorithms have been developed to assist in this identification. However, in this case, a pattern was identified solely through observation.

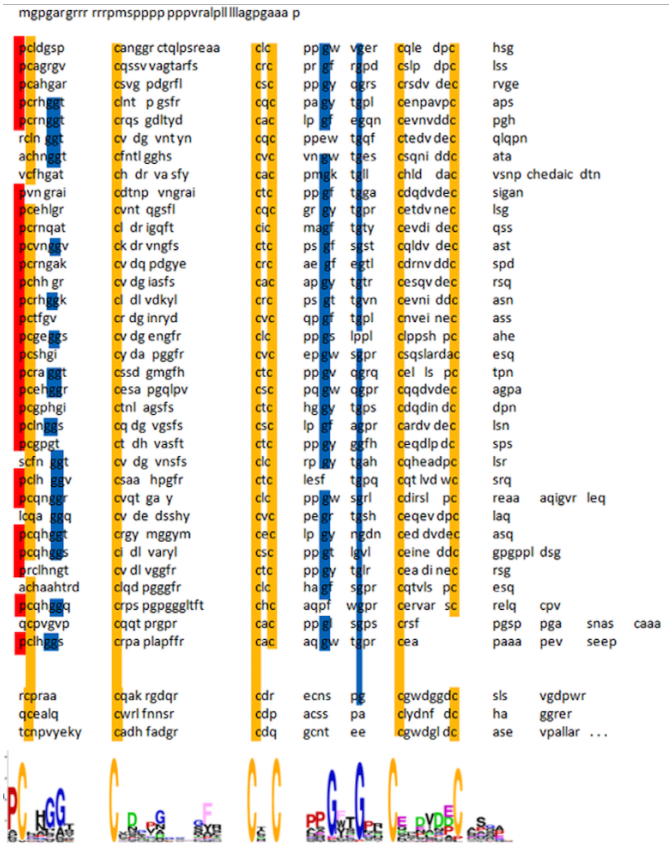

Figure 1. Repeating pattern of the Notch3 protein sequence, divided into repeats, based on the highlighted conserved residues, followed by a Weblogo representing the repeating motif.

The whole Notch3 protein sequences as retrieved from the NCBI protein database is comprised of 2321 amino acids. However, the repeating pattern appears among 1461 amino acids (Figure 1). The latter sequence was divided into 34 most similar repeats and another 3 that demonstrate only partial similarity, as shown. From the last repeat to the end, no pattern is observed.

The alignment was based on the residues highlighted on the sequence. The repeating motif is obvious. The conserved positions which supported the forming of the pattern are apparent on the logo following the sequence (Figure 1).

After having mapped all the point mutations on the sequence of Notch3, as they have been reported from cases examined by Prof. Baumann, the resulting secondary structure is shown in Figure 2. The selected residues represent the mutation positions. The repeating pattern of the mutation positions is obvious in this representation as well. The red arrows indicate the 3 disulphide bonds that are followed by 3 hydrogen bonds. The pattern is repeated 3 times.

As the mutations appear at specific locations on the secondary structure, it is assumable that these locations serve specific purposes, structure-wise, for the Notch3 protein function.

Description of Pattern

As shown in Figure 3A (color distribution of residues based on classification according to structure and chemical characteristics) and 3B (color distribution of residues based on side-chains, charge and hydrophobicity), the repeating pattern starts with a highly conserved proline (pos. 1) which appears on a coil-coil conformation on the secondary structure, followed by a cysteine (pos. 2) almost conserved. Cys residues are responsible for the formation of disulphide bonds, which affect the proper folding and biological function of a protein. The high appearance of mutations on Cys residues, which are mostly conserved, in Notch3 sequence accounts for misfolding and loss of function that may lead to CADASIL disease. The succession of two contiguous higly conserved Gly residues (pos. 5 and 6) constitutes a flexible linker between the two conserved Cys residues (pos. 2 and 8) providing protein stability at that point (Robinson & Sauer 1998, Yan & Sun 1997). The position of the second Gly (pos. 6) seems to be an aliphatic conserved residue, as it is mainly either Gly or Ala. However, Ala does not appear so frequently due to its bulky side chain. On positions 13-17, although not higly conserved, Gly residues are mostly observed, forming a flexible linker once again. The appearance of aromatic residues (F, Y, W) (pos. 18 and 26) right after Gly (pos. 17 and 25) on the strand of the beta sheet on the secondary structure is also noticable. Aromatic residues provide favorable interaction energy through cross-strand pairing with glycine in order to reduce glycine's destabilizing effect (Merkel & Regan 1998). These positions appear conserved by hydrophobic residues (F, Y, A, W, Y) as

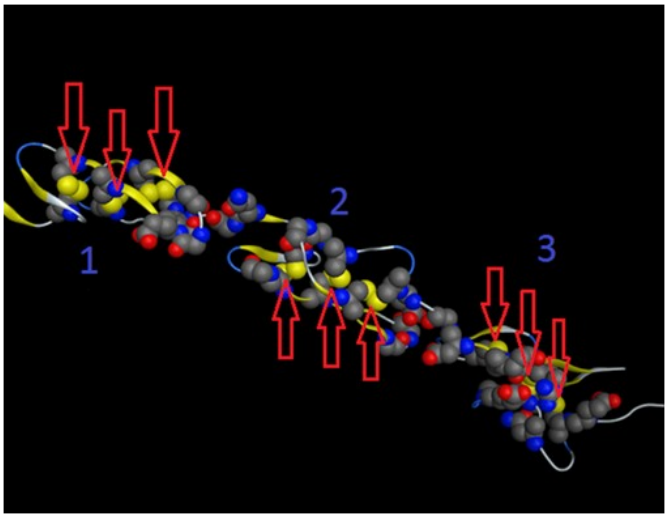

Figure 2. 3D structure of Notch3 protein. The selected residues represent all the point mutations reported by Prof. Baumann. The pattern is obvious in the secondary structure and is repeated 3 times (number above).

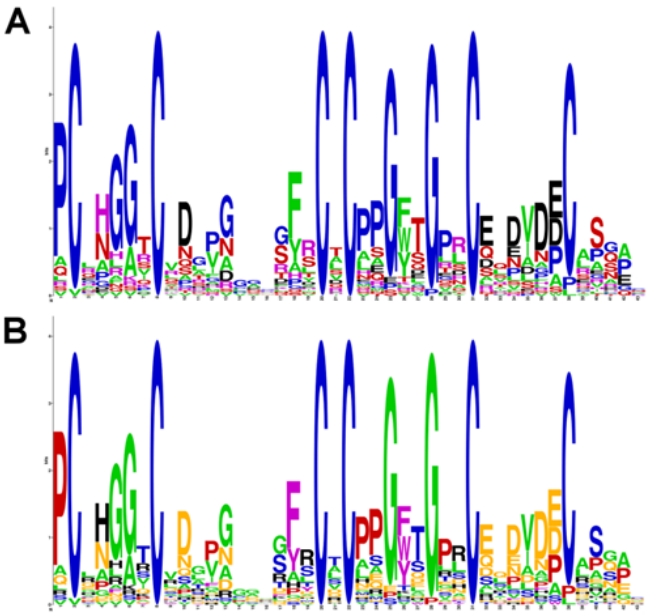

Figure 3. Weblogo representing conserved Notch3 protein repeats. **(A)** Color distribution based on structure and chemical characteristics of residues (green: aliphatic, blue: hydroxyl or sulfur-containing, red: cyclic, black: basic, purple: aromatic, yellow: acidic & their amide). **(B)** Color distribution based on side-chains, charge and hydrophobicity of residues (purple: positively charged, black: negatively charged, red: uncharged, blue: special cases, green: hydrophobic).

well. The Cys-X-Cys motif (pos. 20-22) that follows is completely conservative and is a strong pattern responsible for catalysis of redox reactions (Zhang *et al.* 2008). The two contiguous prolines right afterwards (pos. 23, 24), due to their cyclic character, seem to aid in the formation of the beta turn on the secondary structure. The amino acid sequence that follows contains two quite conservative glycines (pos. 25 and 28), which substitute the smallest amino acid with the simplest side chain of a hydrogen molecule. These two glycines are separated by aromatic residues (F, Y, W) (pos. 26) that have bulky side chains and others with hydroxyl-containing or acidic side chains (T, S, Q, D, E) (pos. 27) that are likely to participate in H-bonding. Position 27 seems to be mostly conserved by uncharged residues (T, S, Q). On positions 31-38, the sequence starts and ends with conserved cysteines with hydrophilic residues (E, D, N) in between, interrupted only by an aliphatic hydrophobic residue (V, I, A) on position 35. This region seems to be exposed to the solvent due to its hydrophilic character. Residues Asp and Glu seem equally replaced by one another at this region, as they are both negatively charged. The high content of Gly and Pro in Notch3, as well as their conservation score in the sequence, serve critical roles for the protein's formation into elastomeric or amyloid fibrils. This finding stands of great significance as amyloid fibrils are inextricably connected with tissue-degenerative diseases, such as CADASIL.

Protein Structure

The Notch extracellular domain is composed of a conserved array of up to 36 epidermal growth factor (EGF)-like repeats, that contribute to ligand interaction, as well as three Lin-12-Notch (LN) repeats, which are juxtamembrane repeats responsible for interactions between the extracellular and intracellular domains. The intracellular domain of Notch comprises of seven ankyrin repeats together with a PEST (proline, glutamine, serine, threonine-rich) and a TAD (transactivation) domain. The C-terminal heterodimerization domain is a hydrophobic region of extracellular Notch, able to create a stable complex with the extracellular domain of transmembrane Notch. The cleavage of the Notch receptor occurs at the S1 site within the trans-Golgi network, during the secretion process. The process of cleavage as well as the resulting structure are vital for Notch function in mammals.

At the surface of the cell, Notch can interact with its ligands, either Delta or Serrate. This interaction sheds the ectodomain and makes the Notch protein vulnerable to further cleavages. Following 3 different cleavages (S2-S4), the intracellular domain of the remaining Notch fragment is released and translocates to the nucleus where it modulates gene expression by acting as a co-activator of the transcription factor suppressor of hairless (Fiuza & Arias 2007).

The negative regulatory region of the Notch receptor, which can be found between the ligand-binding and transmembrane domains, is essential for protecting the Notch receptor against cleavage, when no ligand is present. This region includes three cysteine-rich LIN-12–Notch repeats and a 'heterodimerization domain', which comes right before the membrane and contains both the S1 and S2 cleavage sites. Several studies, over a long time, drove to the conclusion that the negative regulatory region functions as the regulatory switch to activate the receptor. Receptors missing the EGF-like repeats remain functionally inactive. However, deleting the LIN12-Notch repeats or point mutations in this area have been reported to result in gain-of function phenotypes as well as ligand-independent metalloprotease cleavage (Gordon *et al.* 2008).

The NOTCH proteins, *i.e.* NOTCH1-4, consist of 34 EGF-like repeats and work as receptor for Jagged and Delta ligands. They can function as both transmembrane receptors and transcription factors. Upon activation and subsequent release of the notch intracellular domain (NICD), they effect cellular differentiation and proliferation, during development and adult life, including homeostasis in adult tissues (Boucher *et al.* 2012, Fiuza & Arias 2007, Gordon *et al.* 2008).

Molecular Modelling of Notch3 3D Structure

In order to understand the consequences for the NOTCH3 protein when introducing amino acid changes, we took a homology modelling approach since no structure has been determined for this protein. Using Modeller, Notch3 was modelled and its *in silico* 3D structure was established (Sali 1995). Moreover, the consequence of the introduced mutations could also be predicted via molecular dynamics simulations using Gromacs (Hess *et al.* 2008). Modeller predicted a set of two anti-parallel beta-sheet network in arranged repeats across the Notch3 protein. One pair is short, made up of 5 to 6 residues, while the other one is longer as it consists of almost 10 to 12 residues. The beta-sheets are held together by beta turns and coil fragments that contain mainly non-hydrophobic amino acids. A set of strategically positioned cysteine residues is responsible for the structural integrity of the Notch3 structure via disulphide bridges (Figure 4).

The induced R107W mutation provoked the partial loss of structure of one of the beta sheets in the long pair antiparallel conformation of Notch3. The position of the 107 residue was most likely crucial for the proper formation of the first antiparallel set of sheet-loop-sheet, as it was modelled to be located right

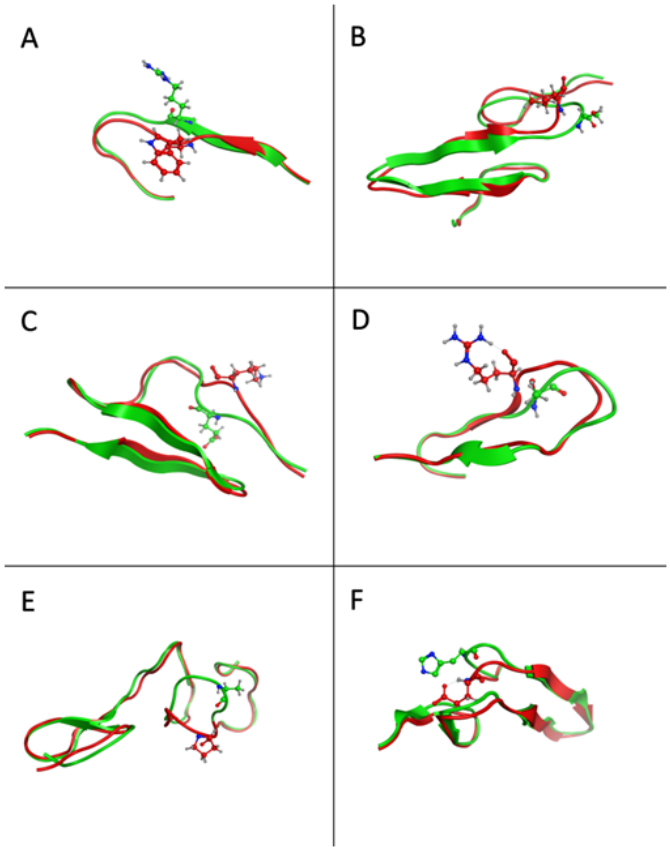

Figure 4. Structural superposition of the wild type and mutant type of Notch3 protein. Color coding is as follows: green for wild type and red for mutant, all shown in ribbon representation. **(A)** Arg107Trp, **(B)** Ser497Leu, **(C)** Glu813Lys, **(D)** Ser978Arg, **(E)** Ala1020Pro, **(F)** His1133Glu.

in the middle of the connecting loop. The bulky and conjugated side chain of the Tryptophan is pushed away from the core of the protein. Therefore it is found to claim the available space in the same plane that the two beta sheets define. As a result, the protein backbone in the proximity of the 107 position is twisted and converted to a coil conformation, having lost its original beta-sheet planarity (Figure 4A).

The S497L mutation was predicted to lead to a partial loss of the first long pair of beta-sheet formations, as it significantly changes the unstructured coil conformation it belongs to, to an almost perpendicular to the beta-sheet plane structure. As a result the network of hydrogen bonds in the nearby anti-parallel beta sheet pair is lost, the gap between the secondary elements becomes slightly bigger and inevitably one of the beta-sheets collapses and loses more than half of its structure (Figure 4B).

The same structural principle applies to the E813K mutation. The negatively charged glutamic acid residue was predicted to actively establish hydrogen bonds to the residues in close proximity that stabilise the nearby pair of beta sheets. The induced E813K mutation introduces a bulky lysine residue in the original glutamic acid position that was modelled to induce breakage of the contact between the two upper beta-sheets. The obvious reason behind this change is attributed to the physicochemical change of the side chain from negatively to positively charged, in the case of the lysine, and to the fact that the lysine is slightly larger in size than the glutamic acid amino acid. It seems the available room is not big enough to accommodate an amino acid larger than glutamic acid (Figure 4C).

The A1020P mutation leads to a significant loss of structure of the nearby beta-sheet arrangement of the Notch3 structure. Even though the 1020P residue is found on a coil-coil structural conformation, the phi/psi dihedral angles of the amino acids involved are very close to a-helical Ramachandran plot regions. Proline cannot establish a hydrogen bond via its backbone amine residue since it has no amide hydrogen. However, even though proline has very poor alpha-helix-forming tendency, it is quite often seen as the first or last residue of an alpha-helix, presumably due to its structural rigidity as its side chain forms a ring. Alanine on the other hand belongs to a group of amino acid residues with low helix forming propensity. In this case, it seems that the induced proline is trying to stabilize an almost alpha-helical local conformation for the small stretch of loop it sits on. Consequently, the structural shift is too great for the nearby beta-sheets to sustain and they collapse, thus losing most of their structure (Figure 4E)

The S978R mutation also leads to loss of beta-sheet structure from Notch3. The Arginine 978 was predicted to be outside of the beta-sheet formation since there is no space available to accommodate it. However, the tiny side chain of the original serine residue was small enough to allow the loop to further bend backwards in favor of the beta-sheet formation. As a result of the induced arginine residue the phi/psi dihedral angle geometry of the nearby residues changes and some residues of the first beta-sheet formation are converted to a coil structure (Figure 4D).

The His1133E mutation is located on a beta-turn conformation that stabilizes the short pair of the anti-parallel beta sheets. The negatively charged glutamic acid is drawn to the inner part of the Notch3 core via newly, randomly established hydrogen bonds to nearby residues. The positive charge of the Histidine's imidazole ring was positioned in a neutral rotamer that allowed the two beta sheets to interact through hydrogen bonding. The tilt of the glutamic acid's side chain and the consecutive rearrangement of the backbone of the protein locally, prohibits the formation of one of the two beta sheets completely. As a result, the Notch3

protein is very much disorganized and locally almost denatured as it completely loses a vital secondary structural element (Figure 4F).

Conclusions

Herein, we presented an *in silico* study of a series of Notch3 mutations that have been recently identified. In most cases these mutations refer to a Cys residue that leads to another unpaired Cys residue. In some cases of this disease, we see accumulation of Granular Osmiophilic Material (GOM), which has been a hallmark for the final diagnosis based on electron microcopy. However, recently some non-Cys mutations have also been identified. These mutations do not follow the characteristic pathology and pattern of the disease. Furthermore, it has been established that some of the non-Cys mutations cause GOM whereas some others do not. The ultimate aim of the *in silico* study was to establish, via the 3D structural analysis, an explanation of what actually happens in these cases. In this direction, exhaustive molecular dynamics simulations showed that these non-Cys mutations trigger significant loss of structure in the Notch3 protein, compared to the wild type. Even though these are mainly point mutations, we have established that the effect of each one of them on the three dimensional structure of the Notch3 protein is significant.

Acknowledgements

This work was partially supported by: (1) The BIOEXPLORE research project. BIOEXPLORE research project falls under the Operational Program "Education and Lifelong Learning" and is co-financed by the European Social Fund (ESF) and National Resources. (2) European Union (European Social Fund - ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: Thales. Investing in knowledge society through the European Social Fund.

Conflicts of interest

The authors declare no conflicts of interest.

References

Balatsos N, Vlachakis D, Chatzigeorgiou V, Manta S, Komiotis D, Vlassi M & Stathopoulos C 2012 Kinetic and *in silico* analysis of the slow-binding inhibition of human poly (A)-specific ribonuclease (PARN) by novel nucleoside analogues. *Biochimie* 94 **1** 214-221

Bellavia D, Checquolo S, Campese AF, Felli MP, Gulino A & Screpanti I 2008 Notch3: from subtle structural differences to functional diversity. *Oncogene* **27** 5092-5098

Bersano A, Ranieri M, Ciammola A, Cinnante C, Lanfranconi S, Dotti MT, Candelise L, Baschirotto C, Ghione I, Ballabio E, Bresolin N & Bassi MT 2012 Considerations on a mutation in the NOTCH3 gene sparing a cysteine residue: a rare polymorphism rather than a CADASIL variant. *Funct Neurol* **27** 247-252

Boucher J, Gridley T & Liaw L 2012 Molecular Pathways of Notch Signaling in Vascular Smooth Muscle Cells. *Front Physiol* **3** 81

Buffon F, Porcher R, Hernandez K, Kurtz A, Pointeau S, Vahedi K, Bousser MG & Chabriat H 2006 Cognitive profile in CADASIL. *J Neurol Neurosurg Psychiatry* **77** 175-180

Chabriat H, Vahedi K, Iba-Zizen MT, Joutel A, Nibbio A, Nagy TG, Krebs MO, Julien J, Dubois B, Ducrocq X *et al.* 1995 Clinical spectrum of CADASIL: a study of 7 families. Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy. *Lancet* **346** 934-939

Consortium TU 2011 Ongoing and future developments at the Universal Protein Resource. *Nucleic Acids Res* **39** D214-D219

Cordle J1, Johnson S, Tay JZ, Roversi P, Wilkin MB, de Madrid BH, Shimizu H, Jensen S, Whiteman P, Jin B, Redfield C, Baron M, Lea SM & Handford PA 2008 A conserved face of the Jagged/Serrate DSL domain is involved in Notch trans-activation and cis-inhibition. *Nat Struct Mol Biol* **15** 849-857

Crooks GE, Hon G, Chandonia JM & Brenner SE 2004 WebLogo: a sequence logo generator. *Genome Res* **14** 1188-1190

Dalkas GA, Vlachakis D, Tsagkrasoulis D, Kastania A & Kossida S 2013 State-of-the-art technology in modern computer-aided drug design. *Brief Bioinform* **14** 745-752

Dichgans M, Mayer M, Uttner I, Brüning R, Müller-Höcker J, Rungger G, Ebke M, Klockgether T & Gasser T 1998 The phenotypic spectrum of CADASIL: clinical findings in 102 cases. *Ann Neurol* **44** 731-739

Ferreira S, Costa C & Oliveira JP 2007 Novel human pathological mutations. Gene symbol: NOTCH3. Disease: cerebral autosomal dominant arteriopathy with subcortical infarcts and leucoencephalopathy (CADASIL). *Hum Genet* **121** 649

Fiuza UM & Arias AM 2007 Cell and molecular biology of Notch. *J Endocrinol* **194** 459-474

Gordon WR, Arnett KL & Blacklow SC 2008 The molecular logic of Notch signaling – a structural and biochemical perspective. *J Cell Sci* **121** 3109-3119

Hess B, Kutzner C, van der Spoel D & Lindahl E 2008 GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J Chem Theor Comput* **4** 435-447

Ishiko A, Shimizu A, Nagata E, Takahashi K, Tabira T & Suzuki N 2006 Notch3 ectodomain is a major component of granular osmiophilic material (GOM) in CADASIL. *Acta Neuropathol* **112** 333-339

Joutel A, Corpechot C, Ducros A, Vahedi K, Chabriat H, Mouton P, Alamowitch S, Domenga V, Cécillion M, Marechal E, Maciazek J, Vayssiere C, Cruaud C, Cabanis EA, Ruchoux MM, Weissenbach J, Bach JF, Bousser MG & Tournier-Lasserve E 1996 Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stroke and dementia. *Nature* **383** 707-710

Joutel A, Dodick DD, Parisi JE, Cecillon M, Tournier-Lasserve E & Bousser MG 2000 De novo mutation in the NOTCH3 gene causing CADASIL. *Ann Neurol* **47** 388-391

Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R & Thornton JM 1996 AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* **8** 477-486

Lewandowska E, Dziewulska D, Parys M & Pasennik E 2011 Ultrastructure of granular osmiophilic material deposits (GOM) in arterioles of CADASIL patients. *Folia Neuropathol* **49** 174-180

Loukatou S, Papageorgiou L, Fakourelis P, Filtisi A, Polychronidou E, Bassis I, Megalooikonomou V, Makalowski W, Vlachakis D & Kossida S 2014 Molecular dynamics simulations through GPU video games technologies. *J Mol Biochem* **3** 64-71

Merkel JS & Regan L 1998 Aromatic rescue of glycine in beta sheets. *Fold Des* **3** 449-455

Monet-Leprêtre M 2009 Distinct phenotypic and functional features of CADASIL mutations in the Notch3 ligand binding domain. *Brain* **132** 1601-1612

Papagelopoulos N, Vlachakis D, Filntisi A, Fakourelis P, Papageorgiou L, Megalooikonomou V & Kossida S 2014 State of the art GPGPU applications in bioinformatics. *Int J Sys Biol Biomed Technol* **2** 24-48

Papageorgiou L, Vlachakis D, Koumandou VL, Papagelopoulos N & Kossida S 2014 Computer-Aided Drug Design and Biological Evaluation of Novel Anti-Greek Goat Encephalitis Agents. *Int J Sys Biol Biomed Technol* **2** 1-16

Pronk S, Páll S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, Shirts MR, Smith JC, Kasson PM, van der Spoel D, Hess B & Lindahl E 2013 GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **29** 845-854

Razvi SS, Davidson R, Bone I & Muir KW 2005 The prevalence of cerebral autosomal dominant arteriopathy with subcortical infarcts and leucoencephalopathy (CADASIL) in the west of Scotland. *J Neurol Neurosurg Psychiatry* **76** 739-341

Robinson CR & Sauer RT 1998 Optimizing the stability of single-chain proteins by linker length and composition mutagenesis. *Proc Natl Acad Sci U S A* **95** 5929-5934

Sali A, Potterton L, Yuan F, van Vlijmen H & Karplus M 1995 Evaluation of comparative protein modeling by MODELLER. *Proteins* **23** 318-326

Scheid R, Heinritz W, Leyhe T, Thal DR, Schober R, Strenge S, von Cramon DY & Froster UG 2008 Cysteine-sparing notch3 mutations: cadasil or cadasil variants? *Neurology* **71** 774-776

Sellis D, Vlachakis D & Vlassi M 2009 Gromita: a fully integrated graphical user interface to gromacs 4. Bioinform Biol Ins **3** 99-102

Sourander P & Walinder J 1977 Hereditary multi-infarct dementia. Morphological and clinical studies of a new disease. *Acta Neuropathol* **39** 247-254

Tatsch K, Koch W, Linke R, Poepperl G, Peters N, Holtmannspoetter M & Dichgans M 2003 Cortical hypometabolism and crossed cerebellar diaschisis suggest subcortically induced disconnection in CADASIL: an 18F-FDG PET study. *J Nucl Med* **44** 862-869

Tikka S, Mykkänen K, Ruchoux MM, Bergholm R, Junna M, Pöyhönen M, Yki-Järvinen H, Joutel A, Viitanen M, Baumann M & Kalimo H 2009 Congruence between NOTCH3 mutations and GOM in 131 CADASIL patients. *Brain* **132** 933-939

Toni-Uebari TK 2013 Cerebral autosomal dominant arteriopathy with subcortical infarcts and leucoencephalopathy (CADASIL): a rare cause of dementia. *BMJ Case Rep* bcr2012007285

Tournier-Lasserve E1, Joutel A, Melki J, Weissenbach J, Lathrop GM, Chabriat H, Mas JL, Cabanis EA, Baudrimont M, Maciazek J, Bach MA & Bousser MG 1993 Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy maps to chromosome 19q12. *Nat Genet* **3** 256-259

Ueda M, Nakaguma R & Ando Y 2009 Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL)]. [Article in Japanese]. *Rinsho Byori* **57** 242-251

Valenti R, Pescini F, Antonini S, Castellini G, Poggesi A, Bianchi S, Inzitari D, Pallanti S & Pantoni L 2011 Major depression and bipolar disorders in CADASIL: a study using the DSM-IV semi-structured interview. *Acta Neurol. Scand* **124** 390-395

Velizarova R, Mourand I, Serafini A, Crespel A & Gelisse P 2011 Focal epilepsy as first symptom in CADASIL. *Seizure* **20** 502-504

Vlachakis D, Argiro A & Kossida S 2013a An update on virology and emerging viral epidemics. *J Mol Biochem* **2** 80-84

Vlachakis D, Bencurova E, Papangelopoulos N & Kossida S 2014 Current State-of-the-Art Molecular Dynamics Methods and Applications. *Adv Protein Chem Struct Biol* **94** 269-313

Vlachakis D, Karozou A & Kossida S 2013b 3D Molecular Modelling Study of the H7N9 RNA-Dependent RNA Polymerase as an Emerging Pharmacological Target. *Influenza Res Treat* **2013** 645348

Vlachakis D, Kontopoulos DG & Kossida S 2013c. Space constrained homology modelling: the paradigm of the RNA-dependent RNA polymerase of dengue (type II) virus. *Comput Math Methods Med* **2013** 108910

Vlachakis D & Kossida S 2013 Molecular modeling and pharmacophore elucidation study of the Classical Swine Fever virus helicase as a promising pharmacological target. *PeerJ* **1** e85

Vlachakis D, Koumandou VL & Kossida S 2013d A holistic evolutionary and structural study of flaviviridae provides insights into the function and inhibition of HCV helicase. *PeerJ* **1** e74

Vlachakis D, Pavlopoulou A, Tsiliki G, Komiotis D, Stathopoulos C, Balatsos NA & Kossida S 2012a An integrated in silico approach to design specific inhibitors targeting human poly(a)-specific ribonuclease. *PLoS One* **7** e51113

Vlachakis D, Tsaniras SC, Feidakis C & Kossida S 2013e Molecular modelling study of the 3D structure of the biglycan core protein, using homology modelling techniques. *J Mol Biochem* **2** 85-93

Vlachakis D, Tsaniras SC & Kossida S 2012b Current viral infections and epidemics of flaviviridae; lots of grief but also some hope. *J Mol Biochem* **1** 144-149

Vlachakis D, Tsiliki G & Kossida S 2013f 3D Molecular Modelling of the Helicase Enzyme of the Endemic, Zoonotic Greek Goat Encephalitis Virus. *Comm Com Inf Sci* **383** 165-171

Vlachakis D, Tsiliki G, Pavlopoulou A, Roubelakis MG, Tsaniras SC & Kossida S 2013g Antiviral Stratagems Against HIV-1 Using RNA Interference (RNAi) Technology. *Evol Bioinform Online* **9** 203-213

Yan BX & Sun YQ 1997 Glycine residues provide flexibility for enzyme active sites. *J Biol Chem* **272** 3190-3194

Zhang T, Zhang H, Chen K, Shen S, Ruan J & Kurgan L 2008 Accurate sequence-based prediction of catalytic residues. *Bioinformatics* **24** 2329-2338