SAPIENZA
UNIVERSITÀ DI ROMA

PhD course in Biochemistry
XXXIV Cycle

Impact of ERK2 missense variants found in cancer: structural, function and stability experimental analysis

PhD Student

Leonore Novak

Coordinators

Prof. Francesco Malatesta

Prof. Stefano Gianni

Tutor

Prof.ssa Roberta Chiaraluce

December 13st, 2021

“Remember to look up at the stars and not down at your feet. Try to make sense of what you see and wonder about what makes the universe exist. Be curious. And however difficult life may seem, there is always something you can do and succeed at.
It matters that you don't just give up.”
Stephen Hawking

Acknowledgement

The search for the most suitable words to thank all those who have accompanied me along this work is exciting and difficult.

Firstly, I would like to express my sincere gratitude to my supervisors Prof.ssa Roberta Chiaraluce and Prof. Valerio Consalvi for believing in me and for instilling in me the qualities of being a good scientist. Your help and guidance over the years has been unmeasurable and without it I would not be where I am today. It has been a pleasure for me work with you.

I am deeply grateful to Dr.ssa Alessandra Pasquo for the knowledge you have passed on. You are my primary re source for getting my science questions answered and I will always be grateful for having the opportunity to learn under your guidance.

Thanks also go out to Prof. Francesco Malatesta for taking care about my professional growth, for his encouragement and helpful advice.

Thank to Prof. Stefan Knapp (SGC, Goethe University of Frankfurt) and his research group to provide me the opportunity to work in his respectable laboratory. Thank to Dr. Apirat Chaikuad for his guidance and for helping me find answers to the scientific questions, without his help I would not have achieved the goals set for me. Thanks to Rezart and Quincin. These two friends formed the core of my research time in the Knapp's group. You are wonderful and generous friends who I admire a lot for your ability to smile despite the situation. I'll never forget the help you gave me in research activities and the wonderful lunches and fun activities we've done together during my period in Frankfurt.

A very special thank you to my friend Maria without his guidance and friendship I would not be who I am today. Working with you was unique and unforgettable and I will never forget it. She has always been my mentor as well as a friend, she has always helped me to believe in me and in my abilities.

A grateful thanks to my colleagues Martina, Maddalena, Daniel, Anna, Valerio, Theo, Roberta, Lucia, Pietro. A good support system is important to surviving and staying sane during PhD. I was lucky to have all of you. You contribute to the following thesis in different ways: giving advices, working hard for helping me, making me smiling.

A special thank to my family for all their love and encouragement. Thanks becouse you supported me in all my choises being always with me and giving me the possibility to reach what I have now. For my parents that teach me what it means believe in something and work hard to reach your goal, and for my sister and her unmeasurable patient and love.

Index

To my family

Chapter 1

1.1 General introduction

Proteins are the most versatile macromolecules in biosystems and serve crucial functions in essentially all biological processes. A reversible process important for the regulation in biological process is protein phosphorylation. Phosphorylation or dephosphorylation can affect the function of a protein in every conceivable way, increasing or suppressing activity, labelling a protein accessible for degradation, allowing it to move from one subcellular compartment to another, or enabling it to interact with or dissociate from other proteins. The abnormal phosphorylation of proteins is known to be a cause of major diseases, such as tumors.

The progression of the neoplastic disease is generally driven by the accumulation of random genetic changes in cells and tissues, which can then develop independence from normal physiological controls due to randomly accumulating mutations. The most common genetic differences in the human genome are single nucleotide polymorphisms (SNPs), which are defined as single nucleotide variations (SNVs) occurring with a frequency of more than 1% in the population. These differences occur on average once every 300–400 base pairs, either in coding or in non-coding regions. SNVs may affect exon splicing or transcription, and are found more frequently than other types of genetic variations, such as differences in copy number, insertions, deletions, duplications, and rearrangements. SNVs in protein-coding regions have received the most attention, in spite of the fact that those regions account for only about 2% of the total human genome. SNVs in the coding region can

be synonymous(sSNVs) if no amino acid change is produced, or non-synonymous(nsSNVs) if the substitution leads to a change in the protein sequence. The nsSNVs can be further divided into two categories: missense mutations, which lead to single amino acid changes, or nonsense mutations, which produce truncated or longer proteins. Missense mutations, that generate protein variants with a single amino acid variation (SAV), are of particular interest in biomedicine, since even just a single amino acid substitution may induce drastic structural alterations, which compromise the protein stability, or may induce crucial structural alterations able to perturb binding interfaces, to the point of impairing the protein function. In particular, this kind of approach seems to be relevant in cancer research considering that several somatic variants resulting from alterations at the amino acid level have been detected in cancer genome for several proteins.

The single amino acid variations detected in the protein kinases and phosphatases, give rise to several of disorders and exert their effects by altering the phosphorylation states of intracellular proteins. The differences between protein kinase variants are at the basis of distinctive traits associated with the susceptibility to specific disease and/or drug response. Many compounds are now marked as ‘specific’ inhibitors of protein kinases, some of which have been used in literally hundreds of papers to study how a protein kinase is implicated in the regulation of a particular cellular event. Nevertheless, very few of these compounds are specific for the protein variants. The information resulting from the analysis of somatic mutations found in cancer tissues can improve the available therapies and create new and more specific ones, the personalized medicine.

1.2 Kinase genome

The human genome, which consists of over 3 billion base pairs, has remained well conserved throughout evolution, in fact it is at least 99.5 % identical between any two humans on the planet [Levy S. et al, 2007]. With the advent of DNA cloning and sequencing in the mid-1970s, it rapidly became clear that a large family of eukaryotic protein kinases exists, and the increasing number of protein kinases led to the speculation that a vertebrate genome might encode as many as 1001 protein kinases [Hunter T., 1987]. The near-completion of the human genome sequence now allows the identification of almost all human protein kinases. The total (518) is about half that predicted 15 years ago, but it is still a strikingly large number, constituting about 1.7% of all human genes [Manning G. et al, 2002]. Manning catalogued the protein kinase complement of the human genome, the "kinome" (Fig. 1), using public and proprietary genomic, complementary DNA, and expressed sequence tag sequences.

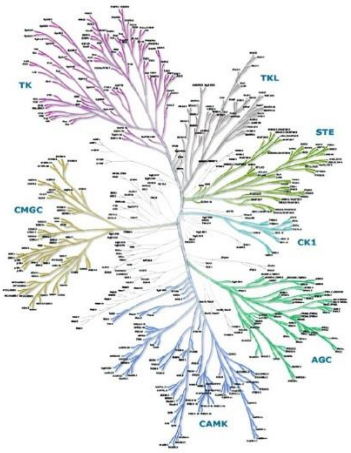

Fig. 1. The human kinase phylogenetic tree [Manning et al, 2002]

This provides a starting point for comprehensive analysis of protein phosphorylation in normal and pathological states, as well as a detailed view of the current state of human genome analysis through a focus on one large gene family.

1.3 Protein kinase

A protein kinase catalyzes the transfer of γ-phosphate from ATP (or GTP) to its protein substrates while a protein phosphatase catalyzes the transfer of the phosphate from a phospho-protein to a water molecule. These enzymes catalyze opposing reactions to modulate the structures and functions of many cellular proteins in prokaryotic and eukaryotic cells. Protein kinases play a key role in almost every aspect of cell biology, and they can modify the function of a protein in almost every conceivable way. Protein phosphorylation can increase or decrease enzyme activity and it can alter other biological activities such as transcription and translation [Roskoski R., 2005]. Conformational changes induced by phosphorylation are highly dependent on the structural context of the phosphorylated protein. Upon phosphorylation, the phosphate group regulates the activity of the protein by creating a network of hydrogen bonds among specific amino acid residues nearby. This network of hydrogen bonds is governed by the three-dimensional structure of the phosphorylated protein and therefore is unique for each protein. [Holt L. J. et al, 2009; Serber Z. et al, 2007]. Conformational changes induced by phosphorylation are highly dependent on the structural context of the phosphorylated protein. Protein kinases employ two types of interactions to recognize their physiological substrates in cells: (i) recognition

of the consensus phosphorylation sequence in the protein substrate by the active site of the protein kinase and (ii) distal interactions between the kinase and the substrate. This latter is mediated by binding of docking motif spatially separated from the phosphorylation site in the substrate and interaction motif or domain located distally from the active site of the kinase [Ubersax J.A.et al, 2007; Turk E.B., 2008]. These interactions contribute to the ability of protein kinases to recognize their protein substrates with specificity. Defining the structural basis of these interactions is expected to benefit identification of potential physiological substrates of protein kinases.

The alteration of phosphorylation by protein kinase is related to many pathological states. The increasing of pathological states has led the studies on the development of new drugs, in particular the protein kinase inhibitors. Protein kinase inhibitors are the front runners in signal transduction therapy because the catalytic sites of enzymes are easier targets for drug design than protein–protein interaction domains.

The protein kinases are involved in many cellular processes; one of the most studied signaling pathways regulated by phosphorylation is Ras-Raf-MEK-ERK cascade. The signaling module of this cascade contains protein-tyrosine, protein-serine/threonine, and dual specificity protein kinases.

1.4 RAS-RAF-MEK-ERK cascade

The Ras/Raf/Mitogen-activated protein kinase/ERK kinase (MEK)/extracellular-signal-regulated kinase (ERK) cascade (Fig. 2) links signals from cell surface receptors to transcription factors, which regulate gene expression. Depending upon the stimulus and cell type, this pathway can transmit signals leading to the prevention or induction of apoptosis or cell

cycle progression. Thus, it is an appropriate pathway to target for therapeutic intervention. This pathway includes multiple members of the kinase and transcription factor families, which can be activated or inactivated by protein phosphorylation. The diversity of signals transduced by this pathway is increased, as different family members heterodimerize to transmit different signals.

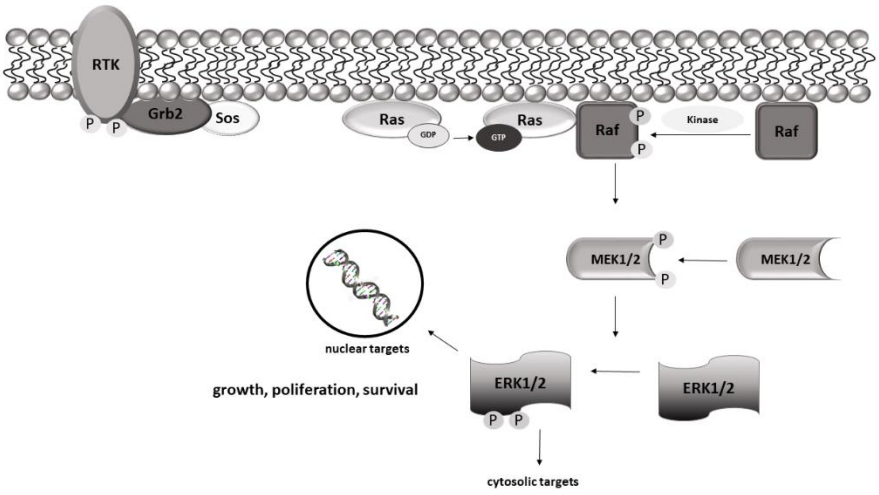

Fig. 2. Ras-Raf-MEK-ERK signal pathways

Ras, an upstream protein of the Raf-MEK-ERK pathway, was the earliest discovered small G protein and product of the *ras* oncogene [Zhou et al., 2016]. It can assume an active GTP-binding conformation and an inactive GDP-binding conformation [Muñoz-Maldonado C. et al, 2019]. The protein can alternate between the two conformations to regulate signal transduction

[Dohlman G. H. et al, 2019]. Ras is activated by many stimulating factors, such as epidermal growth factor, tumour necrosis factor, activators of protein kinase C (PKC) and Src family members [Terrel M. E. et al, 2018]. When an extracellular signal binds to the receptor, a connector molecule, growth factor receptor-binding protein 2 (Grb2), binds to the activated receptor and interacts with the proline-rich sequence at the C-terminus of sevenless (SOS) to form the receptor-Grb2-SOS complex. Binding of SOS to the Y phosphorylation site on the receptor or receptor substrate protein leads to translocation of cytoplasmic SOS to the membrane, resulting in a high concentration of SOS near Ras [Bandaru P. et al, 2019]. SOS and Ras-GDP promote the replacement of GDP with GTP in Ras, thereby activating Ras to initiate the Ras pathway [Simanshu K. D. et al, 2017].

The Raf protein kinase is a protein encoded by the *raf* [Terrel M. E. et al, 2018]. Raf exhibits serine/threonine protein kinase activity after binding to Ras. Raf plays an important role in the Ras/Raf/MEK/ERK cell proliferation signalling pathway [Stokoe et al, 1997]. Activated Raf-1 (due to phosphorylation on residues S338 and S339 by PAK1) continues to activate downstream MEK and ERK, and finally delivers cell proliferation and differentiation signals to the nucleus by regulating the activity of various transcriptional regulators, which regulate gene expression [Vandamme D. et al, 2017].

When Raf is activated, its C-terminal catalytic domain can interact with MEK. The two MEK subtypes, MEK1 and MEK2, have molecular weights of 44 and 45 kDa, respectively [Manning G. et al, 2002]. MEK is a rare dual-specificity kinase that activates ERK by phosphorylating the Y and T regulatory sites, which permits their detachment from the anchoring proteins. [Muta Y. et al, 2019]. Cytoplasmic ERK can phosphorylate a series of other

protein kinases upstream of the ERK pathway, such as SOS, Raf-1 and MEK in a negative feedback regulatory manner [Guo Y.J. et al, 2020]. How MEK has both Y and T specific phosphorylation activity is unclear, but it has important physiological significance, as the ERK signalling pathway is in a central position in the cell signal transduction network, and any errors in activation can profoundly influence cellular processes [Guo Y. J. et al, 2020]. The ERK signaling pathway is in a central position in the cell signal transduction network and dysfunction in the activation can profoundly influence cellular processes. This recognition and activation mechanism that confers double specificity greatly improves the accuracy of signal transduction and prevents errors in ERK activation [Jones J.C. et al, 2017].

The detachment from the anchoring proteins exposes ERK to additional phosphorylation on two serine residues (SPS motif) within the nuclear translocation signal (NTS) of the kinases. This additional phosphorylation allows ERK to interact with importin7, which consequently promotes their translocation to the nucleus [Berti D.A. et al, 2016].

In the nucleus, activated ERK phosphorylates the nuclear transcription factors such as proto-oncogene c-Fos, proto-oncogene c-Jun, ETS domain-containing protein Elk-1, proto-oncogene c-Myc and cyclic AMP-dependent transcription factor ATF2[Roskoski R.,2012].

1.5 ERK 1/2

ERK1 and ERK2 (ERK1/2) are proline-directed kinases that preferentially catalyze the phosphorylation of substrates containing a PXxxS/TP sequence. The regulatory dephosphorylation of ERK1/2 is mediated by protein-tyrosine

specific phosphatases, protein-serine/threonine phosphatases, and dual specificity phosphatases. The combination of kinases and phosphatases makes the overall process reversible. The ERK1/2 catalyzed phosphorylation of nuclear transcription factors such as Ets, Elk, and c-Fos represents an important function and requires the translocation of ERK1/2 into the nucleus by active and passive processes involving the nuclear pore. These transcription factors participate in the immediate early gene response.

ERK1 and ERK2 are very similar in sequence (84%) and usually considered to be functionally redundant, although some differences in their substrate specificity have been reported [Lloyd A.C. et al, 2006; Yoon S. et al, 2009]. Human ERK2 consists of 360 amino acid residues and human ERK1 consists of 379 amino acid residues (Fig. 3) [Roskoski R.,2012].

The two ERK proteins are coexpressed in most tissues but stark differences in their relative abundance were the first possible hints of a differential function. Recent studies [Vantaggiato C. et al, 2006], present compelling evidence that the ERK kinases are not functionally redundant but might have very different roles. These studies showed that ERK2 knockout mice die early in development, suggesting that ERK1 cannot compensate for ERK2 [Yao Y. et al, 2003]. On the contrarily, ERK1-deficient mice are viablewith only minor defects, without any detectable compensatory upregulation of ERK2 level [Yoon S. et al, 2009].

Evidence is mounting that ERK1 and ERK2 have distinct functions. Future studies need to take these findings on board and assess how the interplay between the two kinases affects the signaling dynamics of this pathway and how this can contribute to the cellular response. Therefore, ERK2 is thought to be more significant in cell development and is an attractive target for cancer drug development.

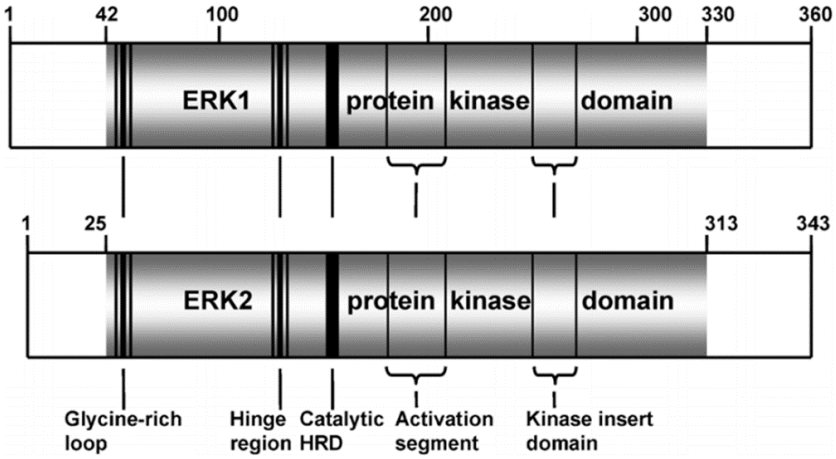

Fig. 3. Architecture of human ERK1 and ERK2 [Roskoski Jr.R.,2012].

1.5.1 ERK2 structure

ERK2, like all protein kinases, have a small amino-terminal lobe and a large carboxyterminal lobe that contain several conserved α-helices and β-strands, first described by Knighton et al. for PKA [Knighton DR. et al, 1991]. The small lobe is dominated by a five-stranded antiparallel β-sheet (β1–β5) [Taylor S.S. et al, 2011]. It also contains an important and conserved αC-helix that exists in active or inactive orientations. The small lobe contains a conserved glycine-rich (GxGxxG) ATP-phosphate-binding loop, sometimes called the A-loop or activation loop, between the β1- and β2-strands (Fig. 5).

The glycine-rich loop, the most flexible part of the N-lobe, helps the position of the ATP β- and γ-phosphates for catalysis. The β1- and β2-strands harbor the adenine component of ATP. The glycine-rich loop is followed by a conserved valine (V39) that makes a hydrophobic contact with the adenine of

ATP (unless otherwise specified, all residue numbers correspond to the human isoforms even when experiments were performed with enzymes from other species). The β3-strand typically contains an AXK sequence, in particular the lysine residue (K54) couples the α- and β-phosphates of ATP to the αC-helix. A conserved glutamate is positioned near the center of the αC-helix (E71). The presence of a salt-bridge between the β3-lysine and the αC-glutamate is a prerequisite for the formation of the activated state and corresponds to the "αC-in" conformation. The αC-in conformation is necessary but not sufficient for the activation of the kinase.

The large C-terminal lobe is mainly α-helical with six conserved segments (αD-αI) [Taylor S.S. et al, 2011]. It also contains four short conserved β-strands (β6-β9) that contain most of the catalytic residues associated with the phosphoryl transfer from ATP to the ERK substrates. The structure of the β-strands is located between those of the αE- and αF-helices, the catalytic core of protein kinases [Knighton DR. et al, 1991]. Of these, the following three amino acids, which define a lysine/aspartate/aspartate (K/D/D) motif, illustrate the catalytic properties of ERK2. An invariant β3-strand lysine (K71 in ERK2) forms salt bridges with the α- and β-phosphates of ATP. The second aspartate of the K/D/D signature, D169, is the first residue of the activation segment.

The activation segments of nearly all protein kinases including ERK2 begins with the sequence DFG and ends with the sequence APE. D169 binds Mg2+ions, which in turn coordinate the α-, β- and γ-phosphates of ATP. The primary structure of the catalytic loop of ERK2, which is positioned near the β6- and β7-strands, contains H147, R148, D149, and K151. The structure of the activation segment is located after that of the catalytic loop and before that of the αF-helix. Functionally important ERK2 residues are listed in Table

1. The large lobe characteristically binds the peptide/protein substrates [Kornev P.A. et al, 2010]. The common name of this site is the CD (common docking) domain. The CD domain is characterized by negatively charged amino acids [D318 and D321 for ERK2], and is located on the opposite side to the active center in the steric structure of the molecules

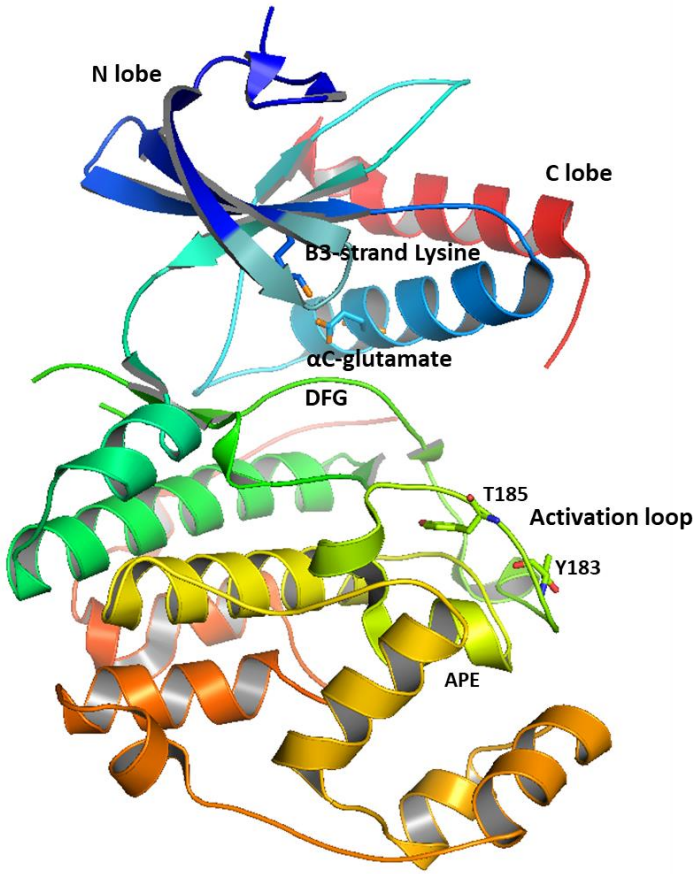

Fig. 5. Structure of ERK2 (PDB: 4zzn) In blue the N lobe, in red the C lobe. In ball and stick blue and cyan respectively are evident the β3-strand Lysine and the αC-glutamate. The activation loop is characterized by the two residues in sticks (Y183 and T185).

Table 1. Important residues in human ERK2[Roskoski R.,2012].

Human ERK2	Human ERK2 residues
Protein kinase domain	42–330
Glycine-rich loop	49–54
The K of K/D/D, or the β3-lysine	71
αC-glutamate	88
Hinge residues	123–126
Gatekeeper residue	Q122
Catalytic HRD	164–166
Catalytic loop lysine	168
DFG of the activation segment	184–186
Activation lip phosphorylation sites	T202, Y204
APE end of the activation segment	212–214
No. of residues	379
Molecular weight (kDa)	43.1
UniProtKB ID	P27361

1.5.2 ERK2 CD docking site

Docking sites are interfaces of protein–protein interaction(PPI), distinct from the catalytic site, which can influence binding interactions according to their position and arrangement. Most biological processes are regulated by a complex network of PPI [Garland W. et al, 2013] that act as a critical regulatory node in many cell-signaling pathways. In some cases, the disruption of PPI due to the presence of mutation in conserved amino acids positioned at interface might lead to both structural and functional loss

[Hardcastle I.R., 2017]. A detailed structural study on residues involved in PPIs is crucial for helping in drug discovery and bioengineering, since the information that can be obtained from this kind of analysis can be helpful in drug design [Larsen C.P. et al, 2005].

In the case of ERKs, docking sites can bind substrates to localize specific Ser/Thr-Pro sequences near the active site to promote their phosphorylation [Sammons R.M. et al, 2019; Sheridan D.L. et al, 2008; Fantz D.A. et al, 2001; Callaway K. et al, 2010; Rainey MA et al, 2005]. ERK2 interacts with its upstream activating enzymes and its inactivating phosphatases through two independent docking domains: the D-site recruitment site (DRS) that interacts with the D-docking domain of substrate and the F-site recruitment site (FRS) that interacts with the F-docking domain of substrates [Robert Roskoski, 2012]. Some substrates possess a D-docking site, some possess an F-docking site, others possess both [Jacobs D. et al, 1999], and some possess neither.

The design of small-molecule inhibitors that block the interaction of the ERK2 recruitment sites and the corresponding substrate docking sites represents a potential strategy for inhibiting the Ras-Raf-MEK-ERK signaling cascade. The D-site of substrates contains both hydrophobic (Φ) and positively charged basic residues with the following canonical sequence: $(R/K)_{2–3}$-$X_{2–6}$-Φ_A-X-Φ_B. This sequence occurs in several transcription factors (Elk1, Lin1, TFII) and other substrates and molecules that interact with ERK2 (STEP, MKP3, PDE4E, MEK1, MNK1, MSK1, and RSK1–3) [Jacobs D. et al, 1999; Sharrocks A.D. et al, 2000]. The ERK2 DRS site is located on the side of the protein that is opposite to the catalytic cleft and consists of a negatively charged component (Φ_{chg}) and a hydrophobic component (Φ_{hyd}). The positively charged basic residues of the substrate D-site bind to the negatively charged ERK2 Φ_{chg} that contains two aspartate residues

(D318/321) previously identified as the common-docking (CD) domain [Tanoue T., 2000]. The Φ_A-X- Φ_B sequence of the ERK2 substrates binds to a nearby hydrophobic patch (Φ_{hyd}) that contains L115, L121, L157, H125, and Y128 (human ERK2 residues).

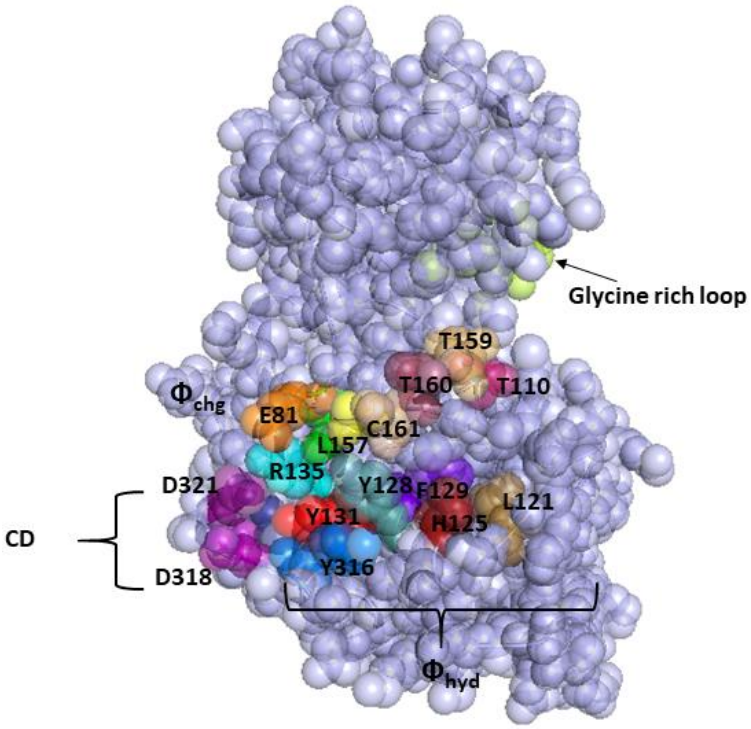

Fig. 6. Model of the human unphosphorylated inactive ERK2 D-site is indicated by the colored space-filling residues that interact with a substrate. Only a small portion of the glycine-rich loop is visible because it is in the "front" of ERK2. CD is the common docking domain. Φ_{chg} corresponds to human ERK2 charged residues, and Φ_{hyd} corresponds to human ERK2 hydrophobic residues. Adapted from the protein data bank file of human ERK2, PDB ID: 4zzn. The residue numbers correspond to those of human ERK2.

The ERK 2 FRS, which binds to a substrate F-site, is located near the activation segment of ERK2. The F-site was first characterized as an FXFP sequence occurring in substrates [Jacobs D. et al, 1999]. This site is conserved in multiple ERK2 interacting proteins including the transcription factors Elk1, c-Fos, Lin1, Sap1, the kinase suppressor of Ras, C-Raf protein kinase, and dual specificity protein phosphatases.

1.5.3 ERK2 inhibitors

The protein kinase inhibitors are divided them into three classes: types I, II, III, IV and V [Dar Arvin C. et al, 2011]. Type I inhibitors bind reversibly to the ATP-binding site of protein kinases and exhibit steady-state enzyme competitive inhibition with respect to ATP; when ATP is bound to the enzyme, type I inhibitors cannot bind. A specific orientation of the αC-helix or the DFG Mg^{2+}binding loop for their effectiveness is not required.

Type II inhibitors occupy the adenine binding site and a region adjacent to it called the back pocket [Zuccotto F. et al, 2010] or hydrophobic pocket II [Zhang J. et al, 2009]. Unlike type I inhibitors, type II inhibitors recognize the conserved DFG motif, which coordinates Mg^{2+}-ATP. They selectively bind to an inactive "DFG-out" conformation, in which the DFG motif is translocated from an active conformation to accommodate a conformation-selective inhibitor. The aspartate from the DFG sequence is rotated nearly180◦ relative to the active state DFG-D in conformation, this rotation leaves empty a hydrophobic site (hydrophobic pocket III), which is available for drug binding. Enzymes that are constitutively active cannot readily assume the DFG-D out conformation.

The type I ½ inhibitors, which is a hybrid of the type I and II classes [Zuccotto F. et al, 2010] bind to the hinge region, the ATP-binding site, the back pocket, and the DFG-aspartate in conformation, not the DFG-out conformation of type II inhibitors. Since type I ½ and II inhibitors occupy part of the ATP site, ATP prevents their binding, and these inhibitors exhibit steady state competitive inhibition with respect to ATP.

Type III inhibitors occupy an allosteric site. These inhibitors block kinase activity while having no effect on ATP binding, and they do not hydrogen bond with the protein kinase hinge region. ATP is unable to displace allosteric inhibitors, and such inhibitors are expected to demonstrate steady-state non-competitive inhibition.

The type IV or irreversible protein kinase inhibitors, characteristically form covalent bonds with their target enzyme [Eglen R. et al, 2011]. Owing to safety and toxicity concerns, covalent irreversible drugs are disfavored as a drug class [Singh J. et al, 2011].

The type V inhibitors bind to two different regions of the protein kinase domain and are therefore bivalent. This type of inhibitors has been classified as bi-substrate inhibitors and interact reversibly with the protein kinase [Roskoski R., 2015].

The first ATP competitive inhibitor was discovered in 2015 by Ohori et al. This molecule, identified as FR180204, is a pyrazolopyridazine derivative, that inhibits the phosphorylation of myelin basic protein catalyzed by human ERK1 (IC50 = 0.51 μM) and ERK2 (IC50 = 0.33 μM). A series of pyrrole compounds are tested to identify the one known as VTX-11e characterized by a marked affinity towards ERK1/2 [Aronov A.M. et al, 2007].

VTX-11e and BVD-523 started the development of many other molecules with similar characteristics, among them the most interesting compound is

SCH772984 [Morris E.J. et al, 2013], which has been shown to inhibit catalytic activity and MEK-dependent phosphorylation [Kidger A.M.et al, 2018]. All the catalytic activity inhibitors analyzed until now are characterized by a portion of N-acyl β-amino alcohol, while SCH772984 is characterized by a different structure in which there is a pyridine-indole portion and an extensive piperazine-phenyl pyrimidine substitution. Indeed, this structure immediately suggested that SCH772984 interacted with a different site compared with a typical type I inhibitors. Several structural studies have shown that SCH772984 induces significant structural modifications between the P-loop and the αC-helix, through the π-stacking interactions with residue Y64 [Morris E.J. et al, 2013]. This conformational variation generates a new binding pocket which is normally not present in either the active or inactive form of the kinase and which in the presence of SCH772984 is occupied by its pyrimidine ring. The DFG motif, on the other hand, is not influenced by the presence of SCH772984 and generally remains in the DFG-in conformation [Chaikuad A. et al, 2014]. Despite the good potency and the discrete selectivity, SCH772984 has a particularly high molecular weight and a poor ligand efficiency which required further efforts to improve the bioavailability of this molecule [Lim J. et al, 2016]. Subsequently, the AstraZeneca researchers have focused on a drug design development to obtain compounds capable of binding effectively to ERK2 in correspondence with the ATP binding pocket on the cysteine residue 166 [Liu Q. et al, 2013]. The lead compound, known as Compound 13, shows good potency and good selectivity [Sipthorp J. et al, 2017]. In fact, although covalent inhibitors could be effective molecules to block downstream ERK signaling, nowadays no promising results have been obtained in this regard.

The only phase 1 clinical trial for covalent inhibitors was initiated by Celgene in 2014 but was stopped because the maximum tolerated dose did not offer a sufficiently encouraging profile to proceed with an additional dosage. This suggests the reactivity of covalent inhibitors and their propensity to give off-target effects [Kidger A. M. et al, 2018].

Besides the compounds directed to the catalytic site, various efforts have been made to develop allosteric inhibitors ERK2, capable of binding to other sites and of interfering with the binding of the protein and its substrates. The first molecule capable of acting as an allosteric inhibitor of ERK was discovered by Shapiro in 2005; in this study an in-silico screening was performed for compounds that could interact with a polar cleft located between the common-docking domain and the E-D motif in the DRS domain in the form of inactive ERK2 [Hancock C.N. et al, 2005]. From this screening, over 80 compounds were selected for subsequent biological assays until the identification of lead Compound 76 (Table 2), which interacts with ERK2 can not interfere with either catalytic activity or phosphorylation [Hancock C.N.et al. 2005]. However, the structural analysis of Compound 76 and other allosteric inhibitors highlighted some critical issues; for example, Compound 76 has an ene-rhodanine core, a motif frequently reported for its marked reactivity.

Table 2. Selective compounds for ERK2

Compunds	Developer	Inhibition type	Clinical progress
FR180204	Atellas Pharmaceuticals	Reversible catalytic / ATP-competitive inhibitor	Pre-clinic
VTX-11e	Vertex pharmaceuticals	Reversible catalytic / ATP-competitive inhibitor	Pre-clinic
BVD-523 (Ulixertinib)	Biomed Valley Discoveries	Reversible catalytic / ATP-competitive inhibitor	Phase 1
SCH772984	Merck	Dual mechanism (catalytic inhibition and MEK-dependent phosphorylation)	Phase 1
AZ13767370 (Compound13)	AstraZeneca	Covalent	Pre-clinic
Compound 76	Shapiro group, University of Maryland	Allosteric	In a clinical study

1.6 Single nucleotide variations

Genetic alterations are diverse and may have several kinds of effects at the phenotypic level. These alterations can cause disease by affecting important biological processes and can affect the react to pathogens, drugs, and vaccines. The most common genetic differences in the human genome are single nucleotide polymorphisms (SNPs), which are defined as single nucleotide variations (SNVs) occurring with a frequency of more than 1% in the population [Collins F.S. et al, 1997, The International HapMap

Consortium, 2003]. SNVs may fall within coding sequences of genes, non-coding regions of genes, or in the intergenic regions (regions between genes). SNVs within a coding sequence do not necessarily change the amino acid sequence of the protein that is produced, due to degeneracy of the genetic code. SNVs in the coding region are of two types, synonymous and nonsynonymous. Synonymous SNVs (sSNVs) do not affect the protein sequence while nonsynonymous SNVs (nsSNVs) change the amino acid sequence of a protein. The nSNVs are of two types: missense and nonsense. A missense mutation is a point mutation in which a single nucleotide change results in a codon that codes for a different amino acid. Another type of nSNVs is a nonsense mutation in which a codon is changed to a premature stop codon that results in truncation of the protein. SNVs in non-coding regions may still affect gene splicing, transcription factor binding, messenger RNA degradation, or the sequence of non-coding RNA. SNVs located within the coding or regulatory regions of genes can cause qualitative and quantitative changes in gene expression, RNA splicing, protein translation, or gene function. As reported by The International Human Genome Sequencing Consortium 2001, only about 5% of the human genome codes to produce proteins and most SNVs are found outside coding sequences [Gan K. A., et al, 2018].

The missense mutations in the coding region generate protein variants with a single amino acid variation (SAV). This SAV are of particular interest in biomedicine, since even just a single amino acid substitution may induce drastic structural alterations, which compromise the protein stability, or may induce crucial structural alterations able to perturb binding interfaces, to the point of impairing the protein function [Zhou T. et al, 2010; Ancien F. et al, 2018]. Usually, change in a single amino acid with similar size and physico-

chemical properties (e.g. substitution from leucine to valine) has a mild effect. Conversely, if the missense mutation disrupts secondary structure elements (e.g. substitution to proline in alpha helix region) such mutation usually may affect whole protein structure and function.

The huge numbers of variants identified over the past twenty years have been collected in several databases, that represent the main source of information for studying the effect of protein variants and understanding the genotype/phenotype relationship [Malhotra et al, 2019]. To maximize the impact of sequence technologies in clinical settings, the scientific community is defining standard protocols and guidelines for discriminating disease-causing variants from nonpathogenic ones [MacArthur D.G. et al, 2014]. Consequently, a large variety of genetic alterations, including nsSNVs, were found to be associated with monogenic and multigenic diseases [Lappalainen T. et al, 2019; Buniello A. et al, 2019], such as type II diabetes mellitus [Flannick J. et al, 2016], acute lymphoblastic leukemia [Liu Y. et al, 2017], or cancer [Bailey H. M. et al, 2018, The ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium, 2020].

Since individuals carrying variants may respond differently to drugs, the structural analysis of nsSNVs may help in understanding how drugs act in individuals with different genetic variants.

1.6.1 Diseases and nsSNVs

A consequence of point mutations in different functional sites can be either the increase or the decrease of protein function. The most dramatic effects have been detected when point mutations occur in the active sites of enzymes or in the binding pockets of receptors [Stevanin G. et al, 2004; Yamada Y. et

al, 2006]. Considering that enzymes catalyze biochemical reactions not only thanks to catalytic residues, but also with several surrounding residues, important for ensuring proper attachment of the substrates and cofactors to the active site cavity (binding sites residues), mutations that occur on the residues located in proximity of the active site can influence the activity [Zhou T. et al, 2010; Takamiya O. et al, 2002; Zhang Z. et al, 2010]. In the human population, 25% of the known nsSNVs significantly affect protein function in vivo [Yue P. et al, 2006]. It has been reported that in most cases nsSNVs lead to loss-of-function (LOF) generally due to perturbation of the active site of the enzyme or of its global structure [Kucukkal T.G. et al, 2015] or to a reduction of the thermodynamic stability of the protein leading to a shift of the folding equilibrium toward the nonfunctional unfolded state, possibly coupled to irreversible aggregation (thermal instability) and/or degradation by cellular quality control [Yue P. et al, 2005; Casadio R. et al, 2011; Shi Z. et al, 2011; Stefl S. et al, 2013; Petukh M. et al, 2015]. Sometimes, but not commonly, nsSNVs can result in a protein that still functions, even if in a "non-canonical" way, as evident in the so-called gain-of-function (GOF) mutants that generally include mutations in proteins such as the Ras GTPase and the epidermal growth factor receptor family of tyrosine kinases, which often drive cancer tumor cell development and proliferation [Schubbert S. et al, 2007; Arteaga C.L. et al, 2014].

The occurrence of a mutational event can generate a protein variant that perturbs conformational constraints of the native protein (e.g., substituting a small side chain residue to a large one and vice versa, resulting in backbone strain or overpacking) or have physicochemical effects (substitutions between hydrophilic residues and hydrophobic residues, burial of charged residues, the disruption of hydrogen bonds, loss of hydrogen bonds, of S–S bonds)

[Shirley B. et al, 1992; Nakao H. et al, 2018; Isom D.G. et al, 2008], leading to considerable alterations of the stability of the native protein [Cristofaro R. et al, 2006; Koukouritaki S.B. et al, 2007; Ode H. et al, 2007].

The ability of proteins to fold correctly into their native and functional form is one of the most fundamental phenomena in nature, and crucial for normal functionality of cells and organisms. [Wang Z. et al, 2001; Ramensky V. et al, 2002; Wang Z. et al, 2003; Capriotti E. et al, 2005; Karchin R. et al, 2005; Ye Y. et al, 2006; Zhang Z. et al, 2010]. It is known that a single nucleotide substitution may result in alteration of stability of the corresponding native protein [Wang Z. et al, 2001; Yue Y. et al, 2005] and that about 80% of missense mutations associated with disease affects the stability of proteins by several kcal·mol^{-1} [Wang Z. et al, 2001]. Generally, a nsSNV can have destabilizing or stabilizing effects: most frequently, missense mutations lead to a destabilization of the protein by making it susceptible to proteolysis or by changing the thermal denaturation temperature. The impact of protein destabilization can be observed in many neurodegenerative diseases, such as Parkinson's disease [Lin W. et al, 2008; Morais V.A. et al, 2009; Nuytemans K. et al, 2010] or in cancer, where destabilizing mutations in the core of the protein may lead to the inactivation of many tumor suppressors [Stehr H. et al, 2011].

Cancer is an important health problem worldwide and it is one of the most frequent causes of death. Cancer is now becoming the first cause of death in developed countries and the second in developing countries [Jemal A. et al, 2011]. The burden of cancer increases rapidly worldwide since the population pressure and the adoption of unhealthy lifestyle, including smoking, drinking, physical inactivity and ‘‘westernized’’ diets [Hua R.H. et al, 2014]. Cancer may indeed be considered an adaptive evolutionary process [Cairns J., 1975;

Greaves M. et al, 2012; Nowell P.C., 1976] and most works correspondingly focus on the study of somatic mutations [Bailey M.H. et al, 2018; Lee B. et al,2018]. As shown by Genome Wide Association Studies (GWAS), many of these cancer-associated genetic variants are not necessarily the cause of the disease, they simply exist into the general picture of the neoplastic disorder and may contribute or not to the evolution of the clinical course of the pathology [Zhang Y. et al, 2019].

The accumulation of mutations in somatic cells represents the most likely event, however, germline mutations are also present in cancer, and they can be associated with cancer susceptibility, which can be passed on to subsequent generations. Public databases (ecg. ClinVar,COSMIC,dpSNP, UniProt humsvar, ICGC Data Portal, OMIM, GDC) report a prevalence of putative pathogenic somatic variants [Milanese J.S et al, 2019], with respect to germline mutations [Chan S.H. et al, 2017,Hu C. et al, 2018].

Even though the exact onset and development of cancer have not been identified, it is determined that genetic susceptibility plays an important role on disease etiology [Lichtenstein P. et al, 2000]. A complete understanding of the mechanisms of the development of cancer is very unlikely to come about in the foreseeable future, making impossible reliance on a single approach to prevent cancer and deaths from the disease. Pharmacological research is of paramount importance from basic research to clinical phase studies. The development of new anticancer drugs can not only improve the efficacy of treatment, but it can reduce the side effects. In cancer research, genetic aspects have become even more important for revealing the molecular basis of the disease that may have a predictive value in cancer development and may contribute to forecast the efficacy of treatment. Treatment throughout history has ranged from removal of organs to targeting specific

markers on cancer cells with molecular drugs. Now, the cancer treatment is even more diverse, other personalized therapies are under development, such as anticancer vaccination and viral gene therapy of cancer by adenovirus particles [Wold W. et al, 2014].

1.6.2 Predicition the effect of nsSNVS

SNVs are variations of single nucleotides that occur between different individuals of the same species. Owing to the redundancy in the genetic code in protein coding genes, these mutations can change the amino acid (nsSNVs), leave the amino acid unchanged (sSNVs) or insert a stop codon. Some methods to predict the effect of nsSNVs and pathogenicity have been developed. Good in-silico methods may help to predict the experimentally observed effects for disease-related variants [Hecht M. et al, 2013]. These methods are mainly based on comparative genomics and/or physical models. These include SIFT [Sim N.L. et al, 2012], Polyphen [Adzhubei I.A. et al, 2010], StSNP [Uzun A. et al, 2007], Bongo [Cheng T.M.K. et al, 2008] and Condel [Ng P.C., 2006; Gonzalez-Perez A. et al, 2011]. According to Venselaar et al, the methods using 3D structure information work better than others [Venselaar H. et al, 2013]. Computational algorithms were implemented to filter the non-significant SNVs from the ones that might produce major disease associated consequences. The phenotypic changes integrated with classical computational SNV prediction techniques will ultimately provide a high accuracy prediction level and thus will help in easy classification of SNVs based on their specific disease-associated consequences. In-silico studies have provided an efficient platform for

evaluation and analysis of genetic mutations for their pathological consequence and in determining their underlying molecular mechanism [Balu K. et al, 2013; Balu K. et al, 2013; Balu K. et al, 2014; Kumar A. et al, 2013].

There are now many databases listing human SNVs and their effects, such as dbSNP (Single Nucleotide Polymorphism database) [Sherry S.T., 2001]; SAAPdb [Hurst J.M. et al, 2009] or SNPdbe [Schaefer C. et al,2012]. Although there are millions of SNVs deposited in public databases, only a small portion of them are functional single nucleotide variants that contribute to disease phenotypes. Thus, prioritizing SNVs based on their phenotypic risks is essential for association studies [Tabor H.K. et al, 2002; Yuan H.Y. et al, 2006]. In addition, the stability effect of a single amino acid mutation should be analyzed at the structural level. Different computational methods have been developed to compare wild type with mutant proteins. In particular, it has been recognized that most disease-causing variants affect the protein thermodynamic stability, expressed as the difference in folding free energy between the native and the denatured state (ΔGf). The experimental determination of the difference in ΔG between the mutant and wild-type proteins (ΔΔGf) constitutes a first and essential analysis for the characterization of the protein variant, using first-principle models or empirical models. Such methods include FoldX [Schymkowitz J. et al, 2005]; PoPMuSiC [Dehouck Y. et al, 2011]; CUPSAT [Parthiban V.et al, 2006]; Rosetta [Leaver-Fay A. et al, 2011], I-Mutant [Capriotti et al, 2005] or INPS (a predictor of the Impact of Non-synonymous-variations on Protein Stability), a new method that computes the ΔΔG values of protein variants without requiring the knowledge of the protein structure [Fariselli P. et al, 2015]. The methods that target understanding the effects of missense mutations on various sequence, structural and functional features have been

explored in the hope of deciphering the phenotype–genotype [Stanley C.M. et al, 2014; Yates C.M. et al, 2013; Kucukkal T.G. et al, 2014]. Computational methods available to predict the effect of single amino acid substitutions on protein stability are based on a single folded structure. However, the native state of a protein is not unique, and it is better represented by an ensemble of its conformers in dynamic equilibrium. The consideration of conformational diversity can improve the discrimination of neutral and disease related protein based on the evaluation of the corresponding Gibbs free energy change [Juritz E. et al, 2012].

The increasing amount of data generated by the several sequencing initiatives [The International Cancer Genome Consortium, 2012; Stratton M.R. et al, 2012] calls for accurate and reliable computational approaches to predict the impact of mutations on the phenotype, and possibly for methods to correlate them with diseases [Casadio et al, 2011]. Despite the availability of high-quality 3D structures (experimental or homology-modeled), of nsSNPs remain a challenging biophysics and bioinformatics problem [Kucukkal et al, 2015].

Although computational methods will not substitute for thorough experimental study, they are useful for obtaining a statistical picture of the extent to which the SNVs influence phenotype and for prioritizing SNVs for experimental study [Hallali-Assani et al, 2009].

Chapter 2

Aim of the study

Available data on polymorphisms in the human genome are expanding rapidly, however knowledge on the possible disease association of nsSNVs and the molecular mechanism of genetic disease is lagging due to the laborious and time-consuming nature of experimental studies.

nsSNVs, or missense mutations, cause amino acid substitutions in the protein product of the gene. nsSNVs are the most common type of genetic variation among humans, and pathogenic nsSNVs account for approximately half of the allelic variants causative of hereditary disease. Amino acid substitutions may have diverse effects on protein structure and function, although some are functionally neutral. Several computational methods for predicting the possible pathogenicity of nsSNVs have been developed. These methods are based on evolutionary information and/or varying structural descriptors of the protein of interest. These methods aim to automating the annotation process of nsSVNs effects and therefore would be very useful for the mutation research community. In-silico studies can give an overall picture of the extent to which the SNPs influence phenotype to give a priority for experimental study [Hallali-Assani et al, 2009].

During my PhD we focused our attention on the characterization of nsSNVs of proteins involved in the regulation of cell metabolism. This thesis reports the comparative analysis of ERK2 protein whose variants are implicated in cancer, one of the most important human disease, an important health problem worldwide and one of the most frequent causes of death.

The variants were selected from somatic mutations databases ATLAS [Cheng P. F. et al, 2015] and COSMIC [Forbes et al., 2010] that store information about genetic variations on the genomic scale and aim to collect mutations in all genes. Other databases, such as dbSNP [Sherry et al., 2001], the Human Genome Variation GenotypetoPhenotype Database (HGVbaseG2P), and the Human Haplotype Map (HapMap) [Frazer et al., 2007], on the other hand, aim to report common allelic variants in the human population that usually have little or no functional consequence, namely nonpathogenic polymorphisms.

We used site-directed mutagenesis to produce recombinant proteins in E. coli cells. After we obtained the variants, we analyzed on protein structure and function, by means of structural stability, binding and activity studies. A detailed study of proteins associated with SNPs and involved in cancer may lead to a better understanding of the biological mechanism for cancer to improve prevention, early detection, and treatment. Our study may provide information also on the rational basis for personalized medicine since nsSNVs have been directly linked to individual susceptibility to develop disease and drugs response.

This comparative and comprehensive study will be useful to understand how missense variants can influence protein properties and which can be the correlation, if any, between the effects induced by single amino acid substitution and cancer.

Chapter 3

Methodology

The aim of our study is to investigate the effect and/or consequences that a nsSNV can induce at the level of the protein. We identified from available databases, such as COSMIC [Forbes S.A. et al, 2010], OMIM [Hamosh A. et al, 2004], Pan-Cancer Atlas [Galka-Marciniak P. et al, 2021] and SNP database, nsSNVs of several proteins found in different cancer tissues. We chose variants of proteins with available PDB to map the single aminoacid substitution on the crystal structure of the protein. We cover all the sequence of ERK2 to analyze the effects of the nsSNVs on the 3D structure of the protein, observing that most of the nsSNVs identified are on the protein surface and only 25% are buried into the internal part of the protein (Fig. 3.1).

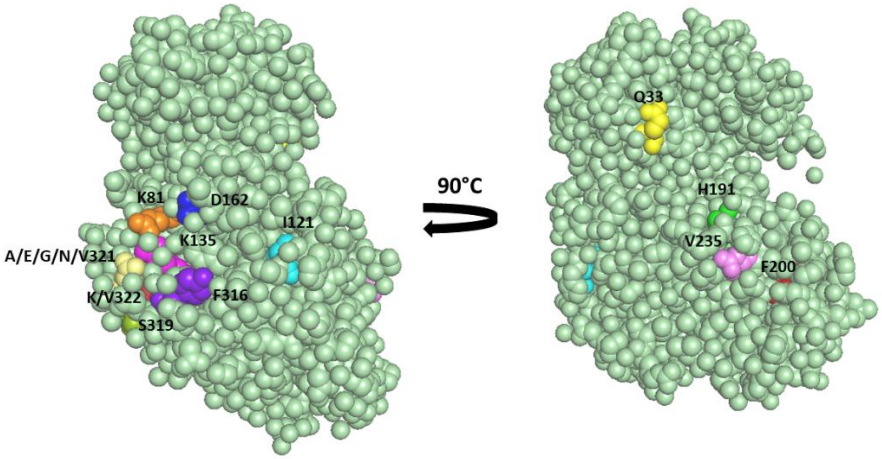

Fig 3.1 Distribution of nsSNVs based on the structural position of the mutated residues.

We designed specific oligonucleotides (Table 1) to obtain the selected variants and we performed site-directed mutagenesis using QuikChange Lightning Site-Directed Mutagenesis (Agilent Technologies, Santa Clara, CA) to introduce the point mutations on wild-type plasmid. The presence of the desired mutations and the absence of unwanted mutations were confirmed by sequence analysis. Once obtained, we expressed the wild-type and the variants as soluble form in *E. coli* and then we carried out a structural characterization by means of circular dichroism and fluorescence spectroscopy. Functional analyses were performed and X-ray crystallography to solve the crystal structure of the variant in the cd docking site to better understand the possible local structural alterations induced by a point mutation. Finally, we performed binding assay using DSF and ITC experiments to better understand the binding properties of the selected variants.

Table 1. The mutagenic oligonucleotides used for site directed mutagenesis of ERK2

Mutant	Primer sequences (5' to 3')
E33Q	FW TCTCGTACATCGGC**CAG**GGCGCCTACGGCAT REV ATGCCGTAGGCGCC**CTG**GCCGATGTACGAGA
E81K	FW GCGCTTCAGACAT**AAA**AACATCATTGGAA REV TTCCAATGATGTT**TTT**ATGTCTGAAGCGC
L121I	FW GAAGACACAACAC**ATC**AGCAATGACCAT REV ATGGTCATTGCT**GAT**GTGTTGTGTCTTC
R135K	FW CTCTACCAGATCCTC**AAA**GGGTTAAAATAT REV ATATTTTAACCC**TTT**GAGGATCTGGTAGAG

D160G	FW AACACCACCTGT**GGT**CTCAAGATCTGTGA REV TCACAGATCTTGAG**ACC**ACAGGTGGTGTT
R191H	FW AATATGTGGCCACA**CAT**TGGTACAGGGCT REV AGCCCTGTACCA**ATG**TGTGGCCACATATT
L200F	FW GGGCTCCAGAAATTATG**TTT**AATTCCAAGGGC REV GCCCTTGGAATT**AAA**CATAATTTCTGGAGCCC
D235V	FW AAGCATTATCTT**GTC**CAGCTGAACCAC REV GTGGTTCAGCTG**GAC**AAGATAATGCTT
Y316F	FW ATATCTGGAGCAG**TTT**TACGACCCGAGTG REV CACTCGGGTCGT**AAA**ACTGCTCCAGATAT
P319S	FW AGCAGTATTACGAC**TCG**AGTGACGAGCCCAT REV ATGGGCTCGTCACT**CGA**GTCGTAATACTGCT
D321A	Fw TATTACGACCCGAGTGCCGAGCCCATCGCCGAA Rev TTCGGCGATGGGCTCGGCACTCGGGTCGTAATA
D321E	Fw TTACGACCCGAGTGAAGAGCCCATCGCCG Rev CGGCGATGGGCTCTTCACTCGGGTCGTAA
D321G	Fw TATTACGACCCGAGTGGCGAGCCCATCGCCGAA Rev TTCGGCGATGGGCTCGCCACTCGGGTCGTAATA
D321N	Fw TATTACGACCCGAGTAACGAGCCCATCGCCGAA Rev TTCGGCGATGGGCTCGTTACTCGGGTCGTAATA
D321V	Fw TATTACGACCCGAGTGTCGAGCCCATCGCCGAA Rev TTCGGCGATGGGCTCGACACTCGGGTCGTAATA
E322K	Fw TACGACCCGAGTGAC**AAG**CCCATCGCCGAA Rev TTCGGCGATGGG**CTT**GTCACTCGGGTCGTA
E322V	Fw TACGACCCGAGTGACAAGCCCATCGCCGAA Rev TTCGGCGATGGGCTTGTCACTCGGGTCGTA

3.1 Protein expression and purification

The wild-type and variants were expressed as N-terminally His-tagged proteins in *E. coli* cells Rosetta with phosphatase under appropriate conditions. The phosphatases expressing strain has been demonstrated to be useful for expression of protein kinases, which would otherwise auto-phosphorylate. The E. coli cells grown at 37°C in LB medium containing kanamycin as an antibiotic at a final concentration of 30 µg/mL until optical density OD_{600} reached 0.6 AU. The protein expression was induced overnight by adding 0.5 mM isopropyl-β-d-thiogalactoside (Sigma-Aldrich, St. Louis, MO) and grown overnight at 20°C with energic shaking. The culture was harvested by centrifugation and resuspended in 40 mL of binding buffer (50 mM Hepes, 500 mM NaCl, 5 mM Imidazole, 5% glycerol *p*H 7.5) containing 0.5 mM Tris(2-carboxyethyl) phosphine in the presence of a cocktail of ethylenediaminetetraacetic acid (EDTA)-free protease inhibitors (Sigma-Aldrich). The cells were sonicated in a Vibracell 75115 sonicator (SONICS, Newtown, CT) with 3 s boots and 9 s pause, on ice and the lysate was cleared by centrifugation. The soluble fraction was applied to a 4-mL prepacked His trap column (GE Healthcare, Chicago, IL) pre-equilibrated with binding buffer. The column was washed with binding buffer to elute weakly bound contaminants and the recombinant protein was eluted by passing over the column binding buffer solutions containing 250 mM imidazole. The eluted protein was concentrated to a final volume of 2.5 mL on an Amicon concentrator Ultra-15 (Millipore, Burlington, MA) and then applied to a PD-10 prepacked column (GE Healthcare) to remove imidazole. The pure fraction in binding buffer was incubated overnight at 4°C with

tobacco etch virus (TEV) protease to cleave the hexahistidine tag. After digestion, the mixture containing TEV protease, the His-tag, and the cleaved protein was applied to a 5-mL prepacked His Trap column (GE Healthcare) previously equilibrated in binding buffer. The flow through containing the protein without His-tag was collected, and purity and size were checked by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) on a pre-casted NuPage 4–12% Bis-Tris polyacrylamide gel (Invitrogen, Carlsbad, CA). The eluted protein was concentrated to a final volume of 2.5 mL on an Amicon concentrator Ultra-15 (Millipore, Burlington, MA) and then applied to a PD-10 prepacked column (GE Healthcare) to change the buffer for further spectroscopic analysis. Protein concentration was determined spectrophotometrically using a molar absorptivity of 44810 M^{-1} cm^{-1} at 280 nm based on a molecular mass of 41.477 kDa.

Furthermore, for the activation of ERK2 we used the plasmid pGEX-KG-MEKR4F (the active mutant of MEK1) plasmid (kindly sent from Professor M. Cobb from Southwestern University, TX, USA) for the co-expression of the wild type and the mutants. ERK2 wild-type and variants will be expressed as N-terminally His-tagged proteins in E. coli cells BL21(DE3)-pLysS under appropriate conditions. The E. Coli cells grown at 37°C in LB medium containing kanamycin (30 µg/mL) and Ampicillin (100µg/mL) until optical density OD_{600} reached 0.6 AU. The protein expression was induced overnight by adding 0.5 mM isopropyl-β-d-thiogalactoside (Sigma-Aldrich, St. Louis, MO) and grown overnight at 20°C with energic shaking. The culture was harvested by centrifugation and resuspended in 40 mL of binding buffer (50 mM Hepes, 500 mM NaCl, 5 mM Imidazole, 5% glycerol, *p*H 7.5) containing 0.5 mM Tris(2-carboxyethyl) phosphine in the presence of a cocktail of ethylenediaminetetraacetic acid (EDTA)-free protease inhibitors

(Sigma-Aldrich). The cells were sonicated in a Vibracell 75115 sonicator (SONICS, Newtown, CT) with 3 s boots and 9 s pause, on ice and the lysate was cleared by centrifugation. For the activated proteins, it was necessary to adjust the protocol for purification by introducing an upstream step: the supernatant was loaded on a DE52 column (GE Healthcare), previously equilibrated with Binding buffer, to remove nucleic acids through ion exchange chromatography that separates molecules based on differences in their net surface charge. The soluble fraction was applied to a 25 mL diethylaminoethyl cellulose (DEAE) pre-equilibrated with binding buffer. The flow through was collected and applied to a 2-mL prepacked His trap column (GE Healthcare, Chicago, IL) pre-equilibrated with binding buffer. The column was washed with binding buffer to elute weakly bound contaminants and the recombinant protein was eluted by passing over the column binding buffer solutions containing 250 mM imidazole. The eluted protein was concentrated to a final volume of 2.5 mL on an Amicon concentrator Ultra-15 (Millipore, Burlington, MA) and we used the dialysis to remove imidazole and to change the buffer for further spectroscopic analysis. Protein concentration was determined spectrophotometrically using a molar absorptivity of 46300 M^{-1} cm^{-1} at 280 nm based on a molecular mass of 43.942 kDa. The protein pure a was checked by two sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) on a pre-casted NuPage 4–12% Bis-Tris polyacrylamide gel (Invitrogen, Carlsbad, CA). In addition to check the ERK1 phosphorylation, we performed western blot analysis with antibodies to doubly phosphorylated ERK2 (Phospho-ERK1/ERK2 (Y183, T185) Antibody (44-680G) in IHC (Invitrogen).

Later, on the same membrane we will perform the stripping protocol with Glycine at *p*H 3 to bind the membrane to the antibody ERK1/ERK2 Antibody

(44-654G) in IF (Invitrogen), to have a double check of correct purification of wild type ERK2 and its variants.

3.2 Spectroscopic measurements

Intrinsic fluorescence emission spectra were recorded from 300–450 nm at 100–130 μg/mL protein concentration containing 20 mM Tris/HCl, *p*H 7.5 0.1 M NaCl and 200 μM DTT with a LS50B spectrofluorimeter (Perkin-Elmer). Fluorescence measurements were carried out using a 1.0 cm path length quartz cuvette at 20°C.

Far-UV (190–250 nm) CD spectra were recorded at a protein concentration ranging over 100–130 μg/mL in 20 mM Tris/HCl, *p*H 7.5 containing 0.1 M NaCl and 200 μM DTT, using 0.1 cm path length quartz cuvette), corresponding to an absorbance of the protein solutions at 280 nm of 0.13 AU.

For Near-UV (240–420 nm) CD spectra, the absorbance of the protein solutions at 280 nm was 1.3 AU, corresponding to a protein concentration ranging over 1.0–1.3 mg/mL in 20 mM Tris/HCl *p*H 7.5, 0.1M NaCl, 1.0 mM DTT, using 1.0 cm path length quartz cuvette. CD measurements were performed in a Jasco-815 spectropolarimeter (Jasco, Easton, MD, Usa) and the results obtained were expressed as the mean residue ellipticity [Θ], assuming a mean residue molecular mass of 110 per amino acid residue.

3.3 Thermal denaturation experiments

ERK2 wild type and variants (100-130 μg/mL) were heated from 20°C to 90°C in 20 mM Tris-HCl, *p*H 7.5, 100 mM NaCl, 0.2 mM DTT, in a 0.1 cm

quartz cuvette with a heating rate of 1 degree x min-1 controlled by a Jasco programmable Peltier element as described in [Lori C. et al, 2013; Petrosino M. et al, 2017]. The dichroic activity at 222 nm and the photomultiplier were continuously monitored in parallel every 0.5°C [Benjwal S. et al, 2006]. The solvent contribution at the various temperatures was taken into consideration for all thermal scans. Melting temperature (*T*m) values were calculated by taking the first derivative of the ellipticity at 222 nm with respect to temperature, as described in Petrosino et al 2017.

3.4 Guanidinium chloride equilibrium unfolding

For equilibrium transition studies, ERK2 wild-type and variants (final concentration 75.0 μg/mL) were incubated at 4 °C at increasing concentrations of GdmCl (0−8 M) in 20 mM Tris/HCl, *p*H 7.5, in the presence of 0.1 M NaCl and 200 μM DTT. After 30 min, equilibrium was reached, and intrinsic fluorescence emission and Far-UV CD spectra (0.2-cm cuvette) were recorded in parallel at 10 °C. To test the reversibility of the unfolding, ERK2 wild-type and variants were unfolded at 10°C in 8.0 M GdmCl at 75.0 μg/mL protein concentration in 20 mM Tris/HCl, *p*H 7.5, in the presence of 2 mM DTT and 0.1 M NaCl. After 5 min, refolding was started by 10-fold dilution of the unfolding mixture at 4 °C into solutions of the same buffer used for unfolding containing decreasing GdmCl concentrations. All denaturation experiments were performed in triplicate.

3.5 Enzyme activity assay and kinetic studies

The biochemical activity of the purified ERK2 wild type and variants has been measured by monitoring the incorporation of phosphate into a peptide substrate. We used a fluorescence-based assay for rapid and sensitive detection of serine/threonine and tyrosine kinase activities, also known as Chelation-Enhanced Fluorescence (ChEF) method. The ChEF sensing mechanism exploits a synthetic α-amino acid with a side chain bearing an 8-hydroxyquinoline derivative (sulfonamido-oxine, Sox) which upon coordination to Mg (II), relays information on the phosphorylation state of proximal serine, threonine or tyrosine residues in peptide- and protein-based kinase substrates. In the absence of phosphorylation, the Sox shows low affinity for Mg (II); upon phosphorylation, Mg (II) affinity is enhanced due to the advantageous chelate effect, involving the Sox and the introduced phosphate group, and fluorescence is turned on [Shults M.D et al, 2006].

ERK2 activity was determined at 30°C with AQT0082 peptide as substrate in a Fluorimeter. The standard reaction mixture contained 54 mM Hepes *p*H 7.5, 10 mM $MgCl_2$, 0.1 M DTT, 0.012% Brij-35, 1% glycerol, 0.2 mg/mL BSA, 5.0 mM MgATP and AQT0082 ranging from 0.069 µM to 40 µM in a final volume of 0.4 mL. The reaction was started by adding different amounts of ERK2 wild type and variants ranging from 8 ng to 24 µg with a dilution obtained in 20 mM HEPES, *p*H 7.5, 0.01% Brij-35, 0.1 mM EGTA, 5% Glycerol, 1 mM DTT, 1 mg/mL Bovine Serum Albumin.

All measurements were performed in triplicate. Kinetic data were analysed using GraphPadPrism 5.04. Results are reported as the mean of three experiments from different protein preparations.

3.6 Temperature dependence of ERK2 activity

The concentrations of peptide used for wild type and variants were choosen according to the K_M values measured for each enzyme. 2 μL of pure enzyme, at 10°C, were added to 0.4 mL of the assay mixture equilibrated at the desired temperature to start the reaction. The final enzyme concentration ranging from 0.5 nM to 1400 nM. The solution was mixed in the thermostated cuvette and the fluorescence at 490 nm was continuously monitored for 5 min. The changes of enzyme activity as a function of temperature were fitted nonlinearly to the Arrhenius equation using GraphPadPrism 5.04 to obtain the activation energies ($E_a^{\ddagger}$) for the catalytic reaction

$$k = Ae\text{-}\ E_a^{\ddagger}/RT$$

where k (s^{-1}) is the rate constant at temperature T (K), A is a reaction specific quantity, R the gas constant (1.987 cal x mol^{-1} x K^{-1}) and $E_a^{\ddagger}$ the activation energy of the reaction, as described in Lori et al. 2013 [Lori C. et al, 2013; Shults M.D et al, 2006].

3.8 Data analysis: quantitative analysis of equilibrium unfolding

The changes in intrinsic fluorescence emission spectra at increasing GdmCl concentrations were quantified as the intensity-averaged emission wavelength, $\bar{\lambda}$, [Royer C. A. et al, 1993] calculated according to

$$\bar{\lambda} = \Sigma(I_i\lambda_i)/\Sigma(I_i) \quad (1)$$

where λi and Ii are the emission wavelength and its corresponding fluorescence intensity at that wavelength, respectively. This quantity is an integral measurement, negligibly influenced by the noise, which reflects changes in the shape and position of the emission spectrum. GdmCl-induced equilibrium unfolding transitions monitored by far-UV CD ellipticities changes were analysed by fitting baseline and transition region data to a two-state linear extrapolation model [Santoro M.M. et al, 1988] according to

$$\Delta G_{unfolding} = \Delta G^{H_2O} + m\,[GdmCl] = -RT\,ln\,K_{unfolding} \quad (2)$$

where $\Delta G_{\text{unfolding}}$ is the free energy change for unfolding for a given denaturant concentration, ΔG^{H_2O} the free energy change for unfolding in the absence of denaturant and *m* a slope term which quantifies the change in $K_{\text{unfolding}}$ per unit concentration of denaturant, R the gas constant, T the temperature and $K_{\text{unfolding}}$ the equilibrium constant for unfolding. The model expresses the signal as a function of denaturant concentration:

$$y_i = \frac{y_N + s_N[X]_i + \left(y_U + s_U[X]_i\right) * \exp\left[\left(-\Delta G^{H_2O} - m[X]_i\right)/RT\right]}{1 + \exp\left[\frac{\left(-\Delta G^{H_2O} - m[X]_i\right)}{RT}\right]} \quad (3)$$

where y_i is the observed signal, y_U and y_N are the baseline intercepts for unfolded and native protein, s_U and s_N are the baseline slopes for the unfolded and native protein, $[X]_i$ the denaturant concentration after the ith addition, ΔG^{H_2O} the extrapolated free energy of unfolding in the absence of denaturant, m the slope of a $\Delta G_{unfolding}$ versus [X] plot. Data were globally fitted with the m values shared between the data sets; all other parameters were not constrained. The denaturant concentration at the midpoint of the transition, $[GdmCl]_{0.5}$, according to equation 2, is calculated as:

$$[GdmCl]_{0.5} = \Delta G^{H_2O}/m \quad (4)$$

All unfolding transition data were fitted by using Graphpad Prism 5.04. Far-UV CD spectra recorded as a function of urea concentration were analyzed by a singular value decomposition algorithm (SVD) using the software MATLAB (Math-Works, South Natick, MA) to remove the high frequency noise and the low frequency random errors and to determine the number of independent components in any given set of spectra. CD spectra in the 213–250 nm or in the 250-320 nm region were placed in a rectangular matrix A of n columns, one column for each spectrum collected at each time. The A matrix is decomposed by SVD into the product of three matrices: $A = U\Sigma V^T$, where U and V are orthogonal matrices and Σ is a diagonal matrix. The U matrix columns contain the basis spectra, and the V matrix columns contain the urea dependence of each basis spectrum. Both U and V columns are arranged in terms of decreasing order of the relative weight of information, as indicated by the magnitude of the singular values in Σ. The diagonal Σ matrix contains

the singular values that quantify the relative importance of each vector in *U* and *V*. The signal-to-noise ratio is very high in the earliest columns of *U* and *V* while the random noise is mainly accumulated in the latest *U* and *V* columns. The wavelength averaged spectral changes induced by increasing denaturant concentrations are represented by the columns of matrix *V*; hence, the plot of the columns of *V* versus the denaturant concentrations provides information about the observed transition.

The ERK2 denaturation processes followed in several cases a different profile due to the formation of an intermediate at low denaturant concentration. The denaturation curve obtained by plotting the fluorescence changes of the ERK2 wild type and variants induced by increasing guanidinium chloride concentrations were fitted to the following equation assuming a three-state model:

$$F = \frac{F_N + \exp\left(m_{I-N}\frac{[GdmCl] - D50_{I-N}}{RT}\right) x \left(F_I + F_U \exp\left(m_{U-I}\frac{[GdmCl] - D50_{U-I}}{RT}\right)\right)}{1 + \exp\left(m_{I-N}\frac{[GdmCl] - D50_{I-N}}{RT}\right) x \left(1 + \exp\left(m_{I-N}\frac{[GdmCl] - D50_{I-N}}{RT}\right)\right)} \quad (5)$$

where F is λ, calculated according to Equation (1); m is a constant that is proportional to the increase in solvent-accessible surface area between the two states involved in the transition; $D50_{I-N}$ and m_{I-N} are the midpoint and m value for the transition between N and I, respectively; and $D50_{U-I}$ and m_{U-I} are the midpoint and m value for the transition between I and U, respectively [Rowling P.J. et al, 2010]. The λ of the intermediate state (I), F_I, is constant, whereas that of the folded state (N) and of the unfolded state (U), F_N and F_U, respectively, has a linear dependence on denaturant concentration:

$$F_N = a_N + b_N [GdmCl] \quad (6)$$

$$F_U = a_U + b_U\,[GdmCl] \quad (7)$$

where a_N and a_U are the baseline intercepts for N and U, and b_N and b_U are the baseline slopes for N and U, respectively. All unfolding transition data were fitted using Graphpad Prism 5.04 (La Jolla, CA, USA).

The three-state model have been used to plot the $[\theta]_{222}$ at increasing denaturant concentrations of ERK2 wild type and some variants

$$y = \frac{y_N + \exp\left(m_{I-N}\frac{[GdmCl]-D50_{I-N}}{RT}\right) x \left(y_I + y_U \exp\left(m_{U-I}\frac{[GdmCl]-D50_{U-I}}{RT}\right)\right)}{1+\exp\left(m_{I-N}\frac{[GdmCl]-D50_{I-N}}{RT}\right) x \left(1+\exp\left(m_{I-N}\frac{[GdmCl]-D50_{I-N}}{RT}\right)\right)} \quad (8)$$

3.9 Western Blot

Gels were briefly rinsed with transfer buffer (Trasfer Buffer 20x (Invitrogen), Antioxidant (Invitrogen) and 10% methanol [v/v]). Protein bands were blotted onto nitrocellulose membranes using a XCell II Blot Module (Novex) at a constant voltage setting of 30 V 250mA for 60 min at 25°C (152.2 mM Trisbase, 1.5 M sodium chloride, 0.1% [vol/vol] polyoxyethylene sorbitan monolaurate [Tween-20], and 5% BSA [w/v], *p*H 7.6). Blots were washed three times, 10 min per wash, the first one in TBS-Tween 0.1%, and the other two wash in TBS. Primary antibodies used in the Western blotting procedures included monoclonal with antibodies to doubly phosphorylated ERK2 (Phospho-ERK1/ERK2 (Y183, T185) Antibody (44-680G) in IHC (Invitrogen) diluted 1:5000 in TBS-Tween. Incubation with primary antibodies was overnight at 25°C. Bound primary antibodies were labelled

with anti-Rabbit IgG (Invitrogeb catalog # 32460), diluted 1:5.000 in TBS, for 60 min at 25°C. Blots were washed in TBS-Tween three times, 10 min per wash, before detection. A chemiluminescent system was used to detect labelled protein bands as described by the supplier (Amersham).

On the same membrane we performed the stripping protocol with Glycine 0.1 M at *p*H 3 for 15 min with agitation. The membrane release of antibody from the antigen was washed 3 times with agitation for 5 minutes each in distillate water and then one time with TBS for 10 min. Membranes were then incubated for 30 min at 25°C in blocking solution (TBST + BSA 5%). Blots were washed three times, 10 min per wash, the first one in TBS-Tween 0.1%, and the other two wash in TBS. After incubating the membrane overnight at 25°C with the ERK1/ERK2 Antibody (44-654G) in IF (Invitrogen), the followed steps were the same used for the previously illustrated procedure.

3.10 Melting Temperature Assay (Differential Scanning Flourimetry (DSF)

The ERK2 wild type and variants were used at the concentration of 2 µM in 10 mM HEPES *p*H 7.5 and 500 mM NaCl. The proteins were mixed with 12 µM inhibitors and SyPRO Orange at 1:1000 dilution. After 10 min incubation at room temperature, the fluorescence signal, which increases upon the temperature-dependent unfolding of ERK2 proteins, was measured using Mx3005p real-time PCR machine (Stratagene, San Diego, CA, USA). The ΔT_m shifts were calculated using protocol described previously Ferorov et al, 2012. The compounds (that have been used) tested were: AZD0364(Item No. 29827), KO-947(Item No. 29213), SCH772984 (Item No. 19166) GDC-0994

(Item No. 21107) LY3214996 (Item No. 27936) Ulixertinib (Item No. 18298) FR180204 (Item No. 15544) Magnolin (Item No. 25123), Pextmetinib (Item No. 26187) and VX-11e (Item No. 19932), purchased from Cayman. and MK-8353 (SCH900353) (Catalog No. S8701), purchased from Selleckchem.

3.11 Isothermal Titration calorimetry

All calorimetric titration experiments were carried out on VP-ITC (MicroCal) at 30 °C. The buffer condition used for ERK2 WT and ERK2 E322K was 20 mM HEPES, *p*H 7.5, 150 mM NaCl and 0.5 mM TCEP. Titration for ERK2 WT and ERK2 E322K was performed by injecting the proteins (170 µM) into a reaction cell containing the inhibitor GDC0994 (20 µM). For the measurements of both peptides one time 4 µL was injected followed by 22 injections of 8 µl with a spacing of 200 s. For the experiment with the peptide pepMEK2 and pepDUSP6 the ERK2 D321N (unphosphorylated and phosphorylated) protein was in the buffer condition of 20 mM HEPES, *p*H 7.5, 150 mM NaCl and 0.5 mM TCEP. For the measurements the peptides were diluted using the same Buffer to a final concentration of 500 µM. The ERK2 wild type and variant was diluted to a final concentration of 90-145 µM. The measurements were performed using an "Affinity ITC" (TA-Instruments) in reversed mode, with a stir rate of 75 rpm and a temperature of 30°C. The protein was placed in the cell and the respective peptide in the syringe. For blank measurements we used the peptides titrated in cell, where was the buffer. For the measurements of both peptides one time 0.5 µL was injected followed by 44 injections of 2 µL with a spacing of 240 s.

The peptides were purchased from GenScript. The pepMEK2 and the pepDUSP6 peptides have the following sequences: MLARRKPVLPALTINP and GIMLRRLQKGNLPVRAL, respectively. They were solubilized in 5% DMSO[w/v] and then diluted in buffer according to the concentrations needed for the experiment. The data were analyzed using the NanoAnalyze Data Analysis software (version 3.10.0; TA-Instruments). The corrected data were fitted to a single binding site model using a nonlinear least-square minimization algorithm, and the binding parameters including reaction enthalpy changes (ΔH), reaction enthalpy changes ($T\Delta S$), equilibrium dissociation constants (K_d), stoichiometry (n) were calculated. The integrated heat of titration was calculated and fitted into an independent binding model using NanoAnalyzer software. The integrated heat of titration was calculated and fitted into an independent binding model using NanoAnalyzer software.

3.11 Crystallization

All crystallization experiments were performed using sitting-drop vapour diffusion method at 20 °C. The variant D321N phosphorylated of ERK2 was buffer exchanged into 20 mM HEPES, *p*H 7.5, 150 mM NaCl, 0.5 mM TCEP, and concentrated up to 9 mg/mL. The variant was incubated with 1 mM of AZD0364 (Cayman Item No. 29827), and crystallized using the reservoir solutions containing 1.6M $MgSO_4$ and 0.1 M MES *p*H 6.0. D321N crystals were small needle clusters and after 1 week harvested.

Chapter 4

Results

The ERK2 (extracellular signal-regulated protein kinases) phosphorylate transcription factors, cytoskeletal proteins, other protein kinases and enzymes mediate key events throughout the cell [Turjanski A.G. et al, 2007; Yoon S. et al, 2006]. ERK2 activity is regulated by phosphorylation on specific threonine (Y183) and tyrosine (T185) residues within their activation loops by MEK. Protein kinases and phosphates work in a balanced state to regulate the function of ERK2 [Hunter T., 1995; Morrison D.K. et al, 2000; Zhou B. et al, 2002; Cheng H.C. et al, 2011; Singh V. et al, 2017]. Phosphorylation is essential for normal cellular processes while, though abnormal phosphorylation is one of the prime causes for alteration of many structural, functional and regulatory proteins in disease conditions, such as cancer.

Interestingly, several ERK2 variants have been identified in tumors as nonsynonymous single nucleotide polymorphisms (nsSNP) that occurr in the DNA coding region and encode a change in the amino acid sequence, leading researchers to some question: “how missense variants in the phosphorylated form and unphosphorylated forms can influence protein properties?

To give an overview of the possible effect or consequences that a single amino acid substitution can have on ERK2, we analyzed 17 nsSNVs of ERK2 protein found in cancer tissues, expressed in the unphosphorylated and phosphorylated forms. Furthermore, ERK2 is an important clinical target with several kinase inhibitors in clinical use or undergoing clinical trials [Carles F. et al, 2018]. A detailed understanding how genetic variations impact the

protein folding, structure and function is required to develop new therapeutic strategies, particularly in the search of small molecules able to selectively interact with the variants, an essential preliminary step to personalized medicine, and help to identify new potential therapeutic targets.

4.1 Experimental analysis of ERK2 nsSNVs

The seventeen human ERK2 variants found in cancer tissues reported in this study were mined from the COSMIC database [Forbes S.A. et al, 2011]. In the Table 4.1 are reported the variants with details about the solvent accessibility, the nucleotide substitution, the pathogenicity, and the cancer tissue.

Table 1 Variants mapped onto the ERK2 structure described on COSMIC database.

	Solvent Accessibility	**Nucleotide substitution**	**Pathogenicity**	**Cancer Tissues**
E33Q	Exposed	G > C	ND	Urinary tract
E81K	Exposed	G > A	0.94	Breast
L121I	Buried	C > A	0.98	Endometrium
R135K	Buried	G > A	0.98	Large Intestine
D162G	Exposed	A > G	1.00	Lung
R191H	Exposed	G > A	0.99	Central nervous system

L200F	Buried	G > T	0.95	Pancreas
D235V	ND	A > T	0.98	Lung
Y316F	Exposed	A > T	0.96	Haematopoietic and lymphoid tissue
P319S	Exposed	C > T	0.98	Skin
D321A	Exposed	A > C	0.98	Liver
D321E	Exposed	C > G	0.85	Urinary tract
D321G	Exposed	A > G	0.97	Lung
D321V	Exposed	A > T	0.98	Skin
D321N	Exposed	G > A	0.98	Upper aerodigestive tract
E322K	Exposed	G > A	0.98	Liver
E322V	Buried	A > T	0.98	Liver

The letter in red indicates the aminoacidic substitution, and the information about nucleotide substitution, pathogenicity and cancer tissues are reported in COSMIC database. The pathogenicity scores for individual mutations from FATHMM-MKL are in the form of a single p-value, ranging from 0 to 1. Scores above 0.5 are deleterious, but to highlight the most significant data in COSMIC, only scores ≥ 0.7 are classified as 'Pathogenic'. Mutations are classed as 'Neutral' if the score is ≤ 0.5.

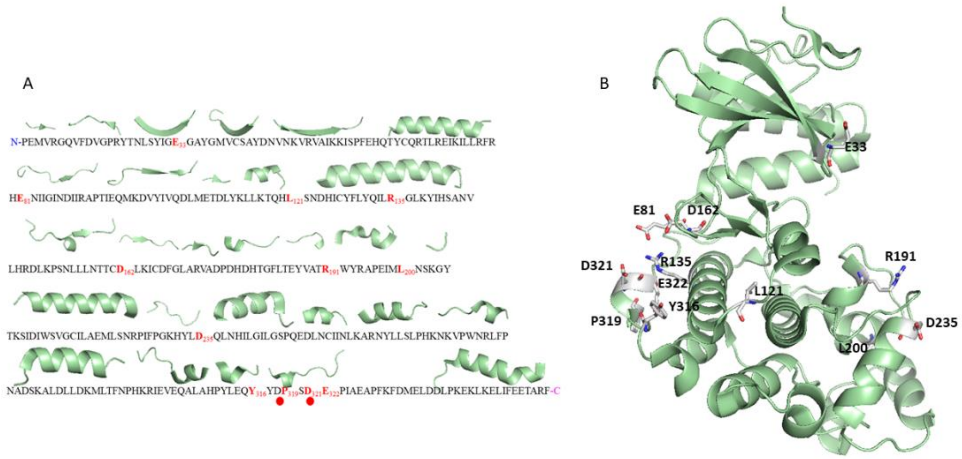

Fig. 1. Amino acid sequence and structure of ERK2. N-terminal domain and C-terminal domain are highlighted in blue and magenta respectively (PDB: 4zzn). (A) Secondary structural elements are shown at the top of the amino acid sequence. Mutated residues are depicted in bold red. The dots under the sequence represent the residues involved in the binding with the substrate (CD-site) (B) Location of the mutations on ERK2 structure (conformation-out). Mutated residues are depicted in stick.

ERK2 contains a common docking domain (CD-site), which includes aspartate residues 318 and 321 (Figure 1A) that are located on the side opposite of the TXY activation loop [Lee T. et al, 2004] and mediates interactions with the substrate. The selected variants involved residues located in the CD-site, in the region that are important for the interaction of the CD-site and in other region of the protein (Fig 1B).

For each of the identified mutant we generated recombinant protein using site-directed mutagenesis. The mutants were expressed in the unphosphorylated (NP-ERK2) and phosphorylated (P-ERK2) form to

understand the structural and biological impact of the protein. To give a clear reading of the results, the variants have been divided into three groups:
1) E81K, R135K, D162G, R191H, Y316F are within or directly adjacent to the acidic region of the CD site, named as “variants interacting with the CD-site”.
2) D321A, D321E, D321G, D321N, D321N, E322K, E322V are located in the CD site, named as “variants in the CD-site”
3) E33Q, L121I, L200F, D325V, P319S are in other regions of the protein, named as “variants in other site of the protein”.
All mutations resulted in soluble recombinant proteins and allowed us to investigate the impact of single amino acid substitution on the ERK2 structure in solution, thermal and thermodynamic stab ility and catalytic activity.

4.2 Spectroscopic properties of ERK2 wild type and variants

The conformation in solution of all ERK2 wild type and variants was studied by circular dichroism (CD) and fluorescence spectroscopy. The near-UV CD spectra of NP-ERK2 and P-ERK2 wild type show a positive contribution at around 290 nm, typical of the tryptophan residue, accompanied by fine structure features at 260–275 nm (Fig. 2A). The two spectra are similar in the tryptophan region, while in the region between 250 and 280 nm the contributions of phenylalanine and tyrosine residues are less pronounced suggesting a conformational change upon phosphorylation.

The fluorescence spectra of the unphosphorylated and phosphorylated ERK2 wild type (Fig. 2B) are centered at the same maximum emission wavelength at around 339 nm, characteristic of tryptophan, with differences in the relative fluorescence emission intensity which is slightly increased in the case of the phosphorylated form.

Far-UV CD spectra of NP-ERK2 and P-ERK2 wild-type show local minima at around 208 and 222 nm and a zero intercept at around 200 nm indicating the major contribution of alpha helical secondary structural elements, slightly influenced by the contribution of β sheet (Fig.2C).

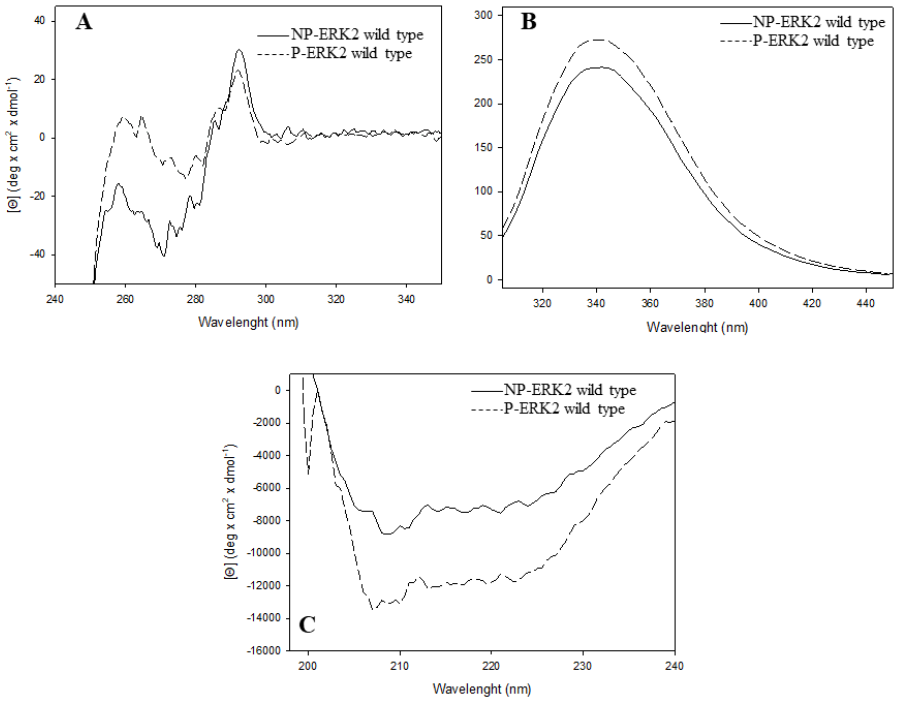

Fig. 2. Spectral properties of NP-ERK2 e and P-ERK2 wild type. (A) Near-UV CD spectra were recorded in a 1.0-cm quartz cuvette at 1.3 mg/mL protein concentration in 20 mM Tris-HCl pH 7.5 containing 1.0 mM DTT and 100 mM NaCl. (B) Intrinsic fluorescence

emission spectra were recorded at 130 μg/mL (0.08 AU280nm, 295 nm excitation wavelength), in 20 mM Tris-HCl, pH 7.5, containing 0.1 M NaCl and 0.2 mM DTT. All spectra were recorded at 20°C. The continuous lines are used for the unphosphorylated form, the medium dashed lines are used for the phosphorylated form.(C) Far-UV CD spectra are recorded in a 0.1-cm quartz cuvette at 130-170 μg/ml in 20 mM Tris-HCl, pH 7.5 containing 0.2 M NaCl and 0.2 mM DTT. All spectra are recorded at 20°C.

The NP-ERK2 variants interacting with the CD site, group I, show a near-UV CD spectra the wild type with the exception of the variant D162G, in which the spectrum shows a lack of tryptophan contribution and an inversion of the molar ellipticity signal (Fig 3A). The spectra of P-ERK2 variants are similar to the wild type (Fig. 3B), except for the variant D162G and E81K. The spectra of P-D162G variant show the same dramatic changes as the NP-D162G, instead for the spectra of P-E81K is observed a decrease of negative ellipticity between 250-270 nm compared with the wild type.

The fluorescence spectra of the NP- and P-ERK2 wild type and variants (Fig. 3C and Fig. 3D) are centered at the same maximum emission wavelength at around 339 nm. In the NP-ERK2 and P-ERK2, significant differences in the relative fluorescence emission intensity have been detected for most of the variants.

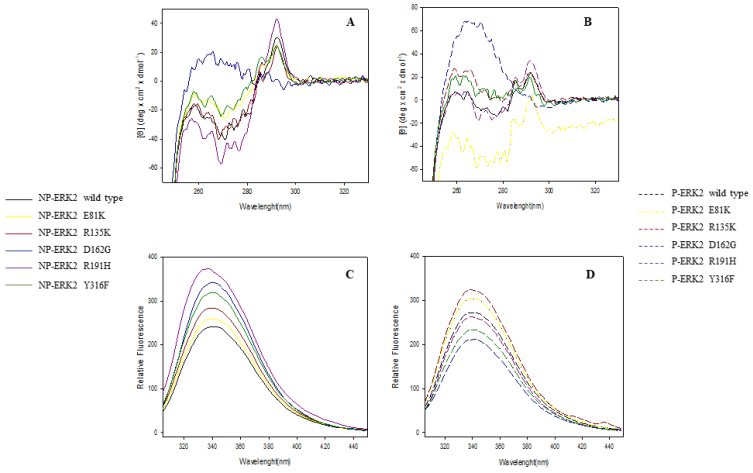

Fig. 3. Spectral properties of NP-ERK2 and P-ERK2 variants interacting with the CD site. All Near-UV CD spectra of NP-ERK2 (the continuous line)(A) and of P-ERK2 variants (the medium dashed lines) (B) are recorded in a 1.0-cm quartz cuvette at 1.3 mg/mL protein concentration in 20 mM Tris-HCl pH 7.5 containing 1.0 mM DTT and 100 mM NaCl All Intrinsic fluorescence emission spectra of NP-ERK2(the continuous line) (C) and of P-ERK2(the medium dashed lines) (D) variants are recorded at 130 μg/mL (0.08 AU280nm, 295 nm excitation wavelength), in 20 mM Tris-HCl, pH 7.5, containing 0.1 M NaCl and 0.2 mM DTT.

The NP-ERK2 variants in the region of CD-site, group II, display a near-UV CD spectra like that of the NP-ERK2 wild type, except for D321G (Fig. 4A) in which the contribution of the tryptophan residue is completely abolished and in the region between 260 and 275 nm are observed a complete inversion of the molar ellipticity signal is observed. Analogously, the P-ERK2 of group II (Fig. 4B) most of the variants show spectra like that of the P-ERK2 wild type with two significant exceptions: D321G, which shows the same spectra observed in the NP-ERK2 D321G and E322K, which shows not well defined contributions in the region 260-275 nm.

The fluorescence spectra (Fig. 4D) of the P-ERK2 D321E variants is centered at 343nm, different from than of the wild type (Fig. 2B). For the NP-ERK2 variants (Fig. 4C), significant differences in the relative fluorescence emission intensity have been detected for all the variants, except for NP-ERK2 D321E and NP-ERK2 D321V.

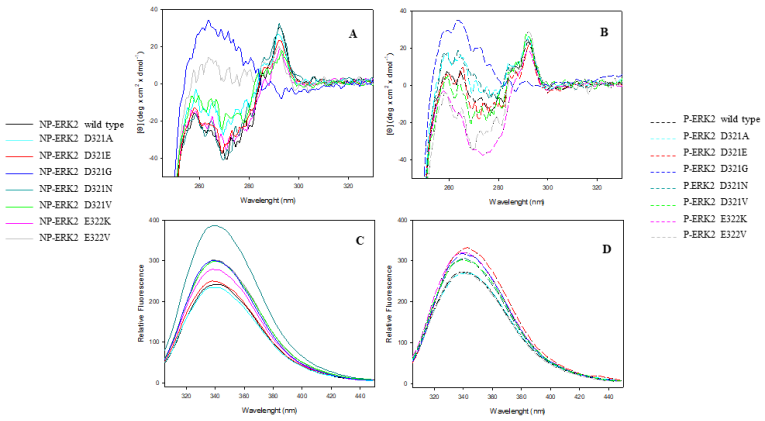

Fig. 4. Spectral properties of NP-ERK2 and P-ERK2 variants in the CD site. All Near-UV CD spectra of NP-ERK2 (the continuous line)(A) and of P-ERK2 variants (the medium dashed lines) (B) are recorded in a 1.0-cm quartz cuvette at 1.3 mg/mL protein concentration in 20 mM Tris-HCl pH 7.5 containing 1.0 mM DTT and 100 mM NaCl All Intrinsic fluorescence emission spectra of NP-ERK2(the continuous line) (C) and of P-ERK2(the medium dashed lines) (D) variants are recorded at 130 μg/mL (0.08 AU280nm, 295 nm excitation wavelength), in 20 mM Tris-HCl, pH 7.5, containing 0.1 M NaCl and 0.2 mM DTT.

The NP-ERK2 variants of group III display spectra similar to that of the NP-ERK2 wild type with the exception of D235V and P319S (Fig. 5A): the contribution of the tryptophan residue is completely abolished and in the region between 260 and 275 nm a complet inversion of the molar ellipticity signal is observed. Most of the P-ERK2 variants (Fig. 5B) show a similar

spectra with respect to the P-ERK2 wild type except for D235V, which shows the same characteristics described for the NP-ERK2 D235V , and the spectrum of L121I, which shows a lack of contribution at 290 nm and a loss in the fine resolution in the 260-275 nm region.

The NP-ERK2 and P-ERK2 variants show differences in the relative fluorescence emission intensity with the exception of E33Q and D235V, that are comparable with the wild type (Fig. 5C and Fig.5D).

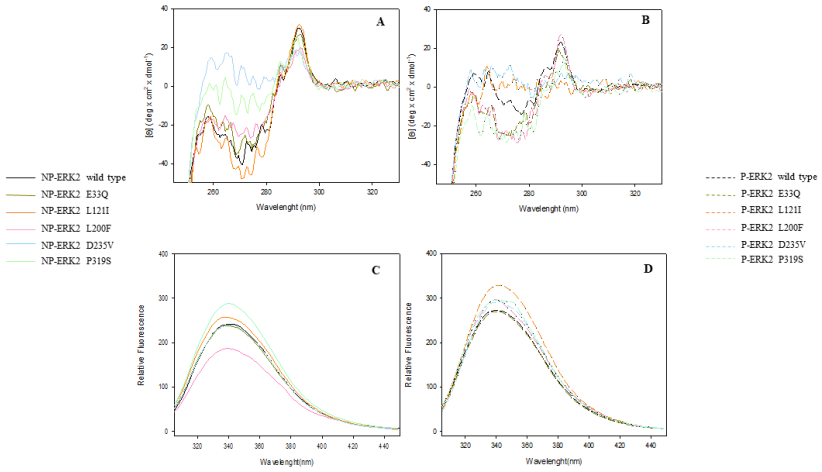

Fig. 5. Spectral properties of NP-ERK2 and P-ERK2 variants in other site of the protein. All Near-UV CD spectra of NP-ERK2 (the continuous line)(A) and of P-ERK2 variants (the medium dashed lines) (B) are recorded in a 1.0-cm quartz cuvette at 1.3 mg/mL protein concentration in 20 mM Tris-HCl pH 7.5 containing 1.0 mM DTT and 100 mM NaCl All Intrinsic fluorescence emission spectra of NP-ERK2(the continuous line) (C) and of P-ERK2(the medium dashed lines) (D) variants are recorded at 130 μg/mL (0.08 AU280nm, 295 nm excitation wavelength), in 20 mM Tris-HCl, pH 7.5, containing 0.1 M NaCl and 0.2 mM DTT.

Far-UV CD spectra of NP-ERK2 and P-ERK2 variants show local minima at around 208 and 222 nm and a zero intercept at around 200 nm indicating the major contribution of alpha helical secondary structural elements, slightly influenced by the contribution of β sheet (Fig.6). As judged by the shape and the intensities of the signals, the secondary structure of the NP-ERK2 variants is native-like (Fig. 6ACE) suggesting that the effect of the point mutations is mainly directed and localized to the mutated residue. The exception is for the spectrum of NP-ERK2 D321G (Fig. 6C), which shows a difference in the intensity and in the local minima around 208 and 222nm, in comparison with that of the wild type.

Furthermore, all the spectra of P-ERK2 variants show a profile similar to that of the wild type, but a divergence in the intensities of the signals, suggesting a different secondary structure rearrangement of the mutants upon phosphorylation.

The molar ellipticity ratio at 222 and at 208 nm ($[\Theta]_{222}/[\Theta]_{208}$) is indicative of interhelical contacts present in helix bundle and coiled coil structures and it is generally used to distinguish between coiled coil helices ($\geq$1.0) and non-interacting helices (0.8-0.9) [Shermann MA et al, 1992; Choy N et al, 2003; Kiss RS et al, 2003]. The ratio of the molar ellipticity at 222 and at 208 nm ($[\Theta]_{222}/[\Theta]_{208}$) observed for the wild type is 0.8. The range is from 0.77 for NP-L121I and P-R191H to 1.16 P-D162G. This variation suggests that the single amino acid substitutions can alter interhelical interactions in solutions (Table 2).

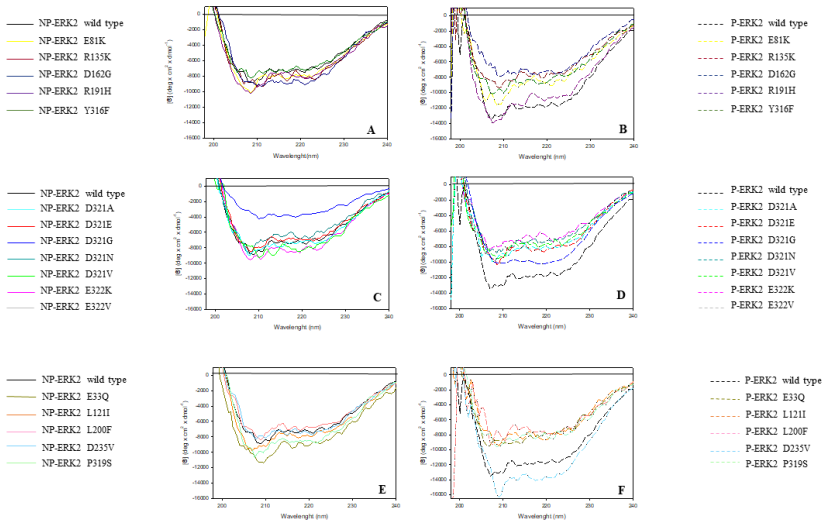

Fig. 6. FAR-UV CD spectra of NP-ERK2 and P-ERK2 wild type and variants. All Far-UV CD spectra are recorded in a 0.1-cm quartz cuvette at 130-170 μg/ml in 20 mM Tris-HCl, pH 7.5 containing 0.2 M NaCl and 0.2 mM DTT. All spectra are recorded at 20°C. (A) NP-ERK2 variants interacting with the CD site (continuous line). (B) P-ERK2 variants interacting with the CD site (the medium dashed lines). (C) NP-ERK2 variants in the CD site (continuous line). (D) P- ERK2 variants in the CD site (the medium dashed lines). (E) NP- ERK2 variants (continuous line). (F) P-ERK2 variants in other (the medium dashed lines).

Table 2 Molar ellipticity ratio at 222 and at 208 nm ([Θ]222/[Θ]208) of NP-ERK2 and P-ERK2 wild type and variants.

ERK2 WT and variants	$[\Theta]_{222/208}$ unphosphorylated form	$[\Theta]_{222/208}$ phosphorylated form
Wild type	0.80	0.90
E33Q	0.80	0.92
E81K	0.80	0.75
L121I	0.77	0.85
R135K	0.90	0.89

D162G	1.01	1.16
R191H	0.83	0.77
L200F	0.83	0.98
D235V	0.88	0.92
Y316F	0.90	0.78
P319S	0.86	0.90
D321A	0.80	0.82
D321E	0.81	0.83
D321G	1.11	0.98
D321N	0.76	0.83
D321V	1.01	0.77
E322K	0.81	0.90
E322V	1.14	0.79

The 222/208 nm ratio for the NP-ERK2 and P-ERK2 wild type and variants is indicative of non-interacting helices (0.8-0.9) and coiled coil helices (>1.0).

4.3 Thermal stability analysis

The thermal stability of NP-ERK2 and P-ERK2 wild-type and variants was investigated in the temperature range between 20°C and 90°C by continuously monitoring the ellipticity changes at 222 nm, where the main amplitude was observed (Fig. 7). The thermal denaturation process is irreversible and occur in an apparent cooperative transition. The parameter chosen to compare the transition curves of the proteins is the melting temperature (T_m) defined as the midpoint of the denaturation process and calculated by plotting the first derivative of the molar ellipticity values as a function of temperature. The

thermal transition curve of NP-ERK2 wild type is less cooperative than the P-ERK2 wild type with a T_m value of 56.0°C and 55.0°C respectively.

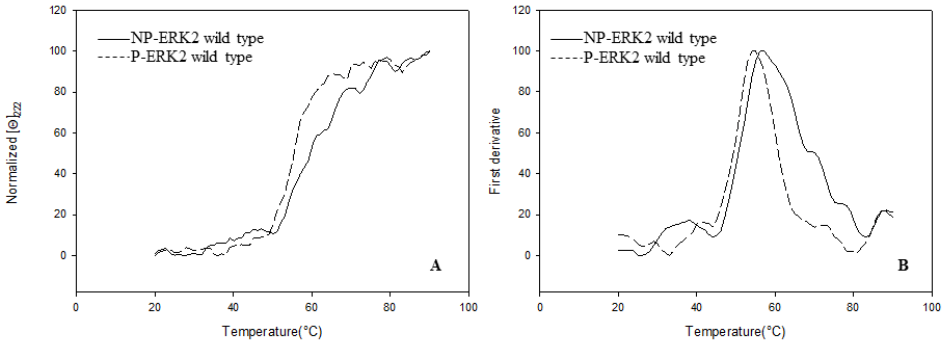

Fig. 7. Thermal unfolding of NP-ERK2 and P-ERK2 wild-type. The ERK2 wild-type (100-130 µg/mL) were heated from 20°C to 90°C in 20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 0.2 mM DTT. The molar ellipticity at 222 nm ($[\Theta_{222}]$) was monitored continuously every 0.5 °C. A) Normalized $[\Theta_{222}]$ B) The first derivative of the same data as in (A).

The variants interacting with the CD site show an apparent cooperative transition (Fig. 8) with a T_m value ranging from 55.0 to 60.0°C. As reported in Table1, a modest increase in T_m value was observed for NP-ERK2 and P-ERK2 R191H and D162G. For D162G, the apparent three-state thermal transitions in the NP-ERK2 suggest the accumulation of unfolding intermediates with two T_m values (Table 3). The first transition occurs with T_m values (T_m^1) at 49.0, the second (T_m^2) at 60.0°C for D162G.

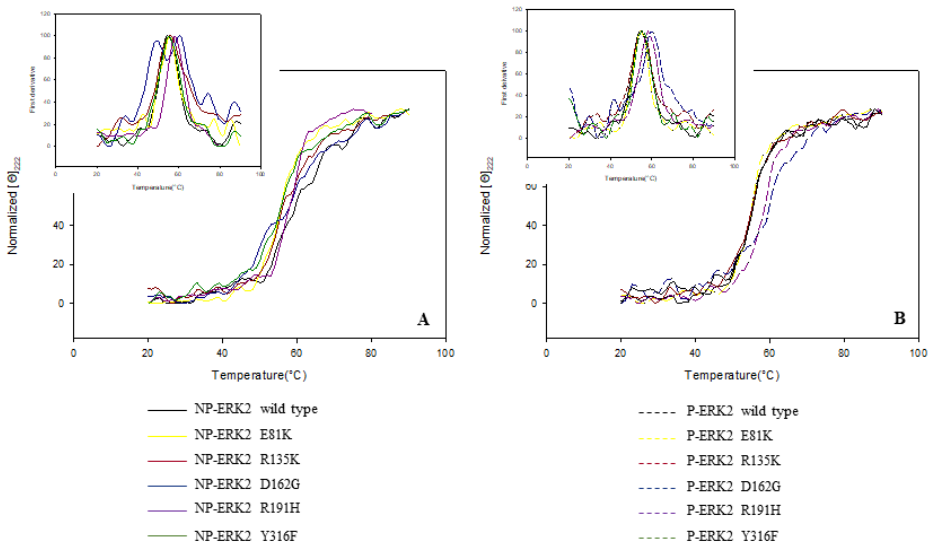

Fig. 8. Thermal unfolding of NP-ERK2 and P-ERK2 variants interacting with the CD site. Normalized [Θ_{222}]; the inset shows the first derivative. ERK2 wild-type and variants (103 μg/ml - 130 μg/ml) were heated from 20°C to 90°C in 20 mM Tris-HCl, pH 7.5 containing 0.2 M NaCl and 0.2 mM DTT. The molar ellipticity at 222 nm ([Θ_{222}]) was monitored continuously every 0.5 °C. (A) Thermal unfolding of NP-ERK2 and P-ERK2 wild type and variants (continuous line) (B) Thermal unfolding of P-ERK2 and P-ERK2 wild type and variants (the medium dashed lines).

Table 3 Melting temperatures for NP-ERK2 and P-ERK2 wild-type and variants interacting with the CD site measured by far-UV CD.

	T_m(°C) unphosphorylated	**T_m(°C) phosphorylated**
Wild type	56.0	55.0
E81K	55.0	54.1

R135K	55.1	55.1
D162G	T_{m1} 49.0 T_{m2} 60.0	60.0
R191H	58.0	58.1
Y316F	55.1	55.1

The temperature-induced changes were followed by monitoring the ellipticity at 222 nm. The T_m values were calculated by taking the first derivative of the ellipticity at 222 nm with respect to temperature.

The variants in the CD site show some relevant differences in thermal unfolding curve (Fig. 9). A modest increase in T_m values is observed for the NP-ERK2 and P-ERK2 D321A, D321N, and E322V variants (Table 4), especially for the variant D321N, whose T_m value is 4 degrees higher than that of the NP-ERK2 wild type. In the case of P-ERK2 D321V the T_m value is 54.5°C three degrees lower than that of NP-ERK2 D321V (T_m 57°C). The NP-ERK2 and P-ERK2 D321E, D321G, E322K variants show T_m values lower than the wild-type, particularly E322K that shows a T_m value four degrees below in the NP-ERK2 and six degrees below in the P-ERK2 compared to that of the wild type.

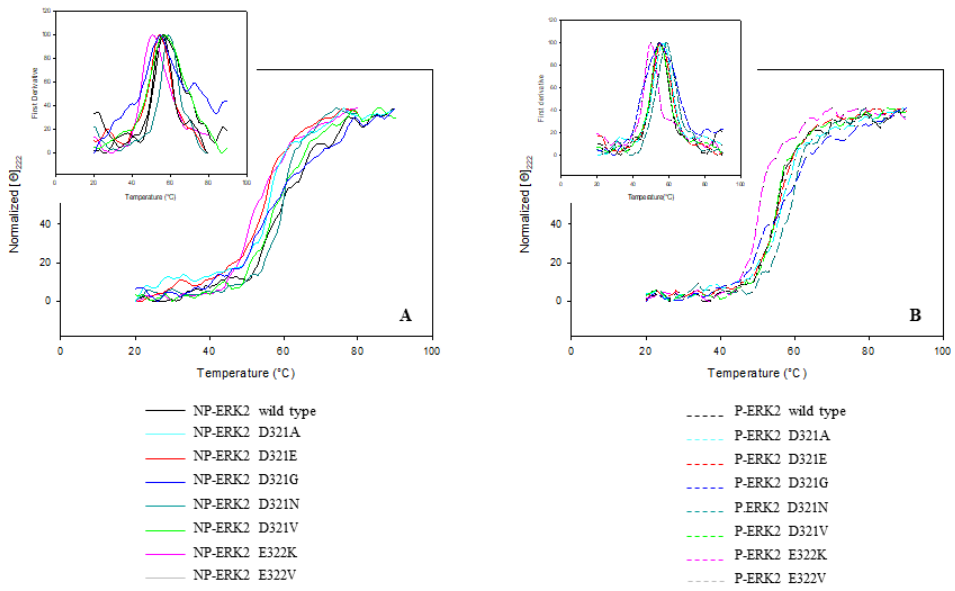

Fig. 9. Thermal unfolding of NP-ERK2 and P-ERK2 wild-type and variants in the CD site. Normalized [Θ_{222}]; the inset shows the first derivative. ERK2 wild-type and variants (103 μg/ml - 130 μg/ml) were heated from 20°C to 95°C in 20 mM Tris-HCl, pH 8.0 containing 0.2 M NaCl and 0.2 mM DTT. The molar ellipticity at 222 nm ([Θ_{222}]) was monitored continuously every 0.5 °C. (A) Thermal unfolding of NP-ERK2 and P-ERK2 wild type and variants (continuous line) (B) Thermal unfolding of P-ERK2 and P-ERK2 wild type and variants (the medium dashed lines).

Table 4 Melting temperatures for NP-ERK2 and P-ERK2 wild-type and variants in the CD site measured by far-UV CD.

	T_m(°C) unphosphorylated	T_m(°C) phosphorylated
Wild type	56.0	55.0
D321A	56.0	57.1
D321E	54.0	55.1

D321G	54.0	55.0
D321N	59.0	58.1
D321V	57.0	54.5
E322K	51.0	50.3
E322V	57.1	56.0

The temperature-induced changes were followed by monitoring the ellipticity at 222 nm. The T_m values were calculated by taking the first derivative of the ellipticity at 222 nm with respect to temperature.

The variants in other site of the protein, group III, display an apparent cooperative transition for NP-ERK2 and P-ERK2 (Fig. 10), with T_m values ranging from 51.0 to 62.0°C (Table 5). In the NP-ERK2 all the variants show a decrease or a similarity in the T_m values compared to that of the wild type. The T_m values of P-ERK2 variants are higher than that of the wild type, only two exceptions, L200F and P319S that show the same T_m value of the wild-type.

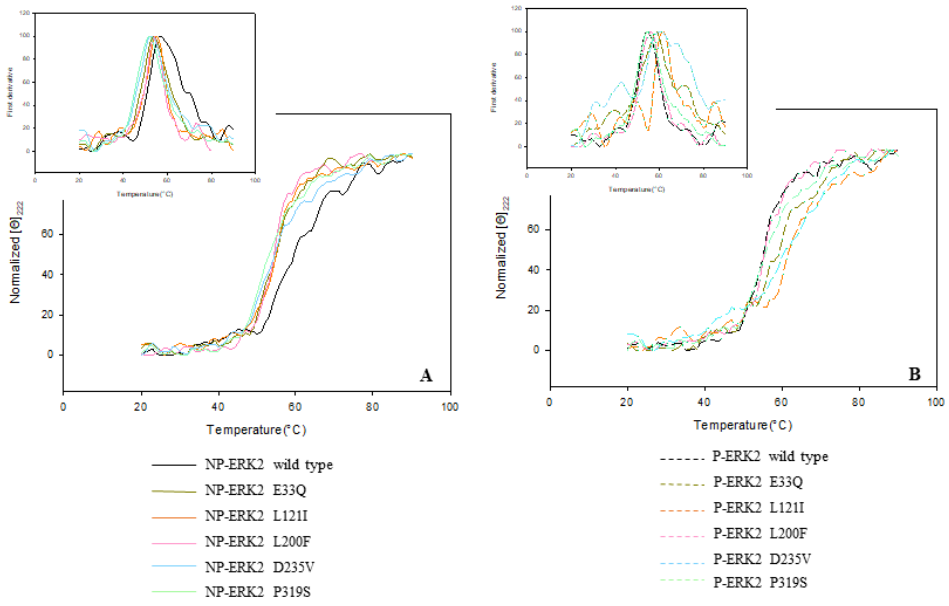

Fig. 11. Thermal unfolding of NP-ERK2 and P-ERK2 wild-type and variants in other site of the protein. Normalized [Θ_{222}]; the inset shows the first derivative. ERK2 wild-type and variants (103 μg/ml - 130 μg/ml) were heated from 20°C to 95°C in 20 mM Tris-HCl, pH 8.0 containing 0.2 M NaCl and 0.2 mM DTT. The molar ellipticity at 222 nm ([Θ_{222}]) was monitored continuously every 0.5 °C. (A) Thermal unfolding of NP-ERK2 and P-ERK2 wild type and variants (continuous line) (B) Thermal unfolding of P-ERK2 and P-ERK2 wild type and variants (the medium dashed lines).

Table 5 Melting temperatures for unfolding equilibrium of NP-ERK2 and P-ERK2 wild-type and variants in other site of the protein measured by far-UV CD.

	T_m(°C) unphosphorylated	T_m(°C) phosphorylated
Wild type	56.0	55.0
E33Q	54.1	59.0
L121I	55.0	62.0
L200F	54.0	56.1
D235V	52.1	61.1
P319S	51.0	55.1

The temperature-induced changes were followed by monitoring the ellipticity at 222 nm. The T_m values were calculated by taking the first derivative of the ellipticity at 222 nm with respect to temperature.

4.4 Thermodynamic unfolding analysis

ERK2 wild-type and variants reversibly unfolds in GdmCl in 20 mM TrisHCl, pH 7.5, containing 0.1 M NaCl and 200 μM DTT. Incubation of ERK2 wild type and variants at increasing GdmCl concentrations at 10°C resulted in a progressive change in 222 nm ellipticity by far-UV CD and in the intrinsic fluorescence emission intensity with a red- shift of the maximal emission wavelength $\overline{\lambda}$. The same samples, during the unfolding transitions, were used to monitor the fluorescence changes to monitor the far-UV CD ellipticity changes, to allow a direct comparison of the data. The fluorescence changes at increasing GdmCl concentration were measured by calculating the

intensity averaged emission wavelength $\bar{\lambda}$, an integral measurement that depends both on the position and the shape of the spectrum.

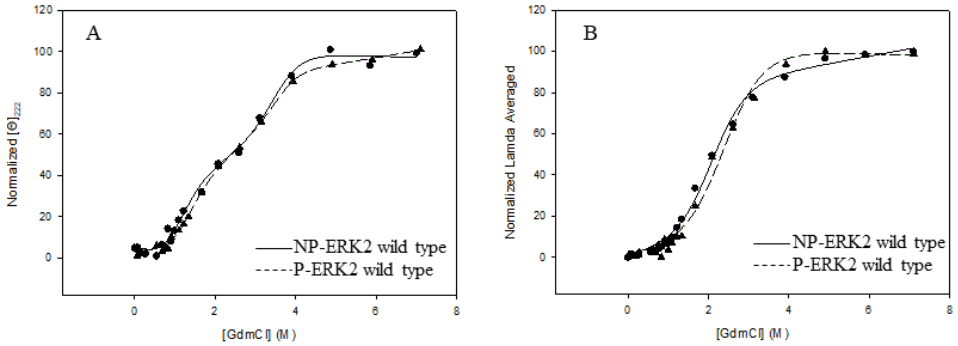

Fig. 12. GdmCl induced equilibrium unfolding of NP-ERK2 and P-ERK2 wild-type. (A) Normalized molar ellipticity at 222 nm ($[\Theta_{222}]$) reported after removal of the high-frequency noise and the low-frequency random error by SVD. The continuous lines represent the nonlinear fitting of the normalized $\bar{\lambda}$ and $[\Theta_{222}]$ data to Eq 3 and Eq 8. The reversibility points (empty circles) are shown, for clarity, only for the wild-type and were not included in the nonlinear regression analysis. (B)Normalized intensity-averaged emission wavelength $\bar{\lambda}$. All spectra were recorded at 10 °C.

The unfolding process of the NP-ERK2 wild type monitored by following the changes in the intrinsic fluorescence follows an apparent two-state transition without any detectable intermediate (Fig. 12). The unfolding transition analysed by the far-UV CD ellipticity change, shows a three-state process due to the formation of an intermediate at low denaturant concentration. For the P-ERK2 wild type the unfolding process followed by

the change in the intrinsic fluorescence and the far-UV CD ellipticity shows a two-state transition(Fig.12).

The effect of increasing GdmCl concentrations (0–8 M) for ERK2 wild type was used for a comparison with the variants. The unfolding process is fully reversible upon dilution of the denaturant also for the variants.

Table 6 reports the variants following the two-state, in white, or three state in grey.

Table 6 ERK2 wild type and variants equilibrium unfolding transition.

	CD ($[\Theta]_{222}$)	Fluorescence (Lambda Averaged)
Wild type unphosphorylated	Three state transition	Two-state transition
Wild type phosphorylated	Two-state transition	Two-state transition
E33Q unphosphorylated	Two-state transition	Two-state transition
E33Q phosphorylated	Three state transition	Two-state transition
E81K unphosphorylated	Three state transition	Three state transition
E81K phosphorylated	Three state transition	Three state transition
L121I unphosphorylated	Two-state transition	Two-state transition
L121I phosphorylated	Two-state transition	Two-state transition
R135K unphosphorylated	Three state transition	Two-state transition
R135K phosphorylated	Two-state transition	Two-state transition
D162G unphosphorylated	Two-state transition	Two-state transition
D162G phosphorylated	Three state transition	Two-state transition
R191H unphosphorylated	Three state transition	Two-state transition
R191H phosphorylated	Two-state transition	Two-state transition
L200F unphosphorylated	Three state transition	Two-state transition
L200F phosphorylated	Two-state transition	Two-state transition

	CD ($[\Theta]_{222}$)	Fluorescence (Lambda Averaged)
D235V unphosphorylated	Two-state transition	Two-state transition
D235V phosphorylated	Two-state transition	Two-state transition
Y316F unphosphorylated	Two-state transition	Two-state transition
Y316F phosphorylated	Two-state transition	Two-state transition
P319S unphosphorylated	Three state transition	Three state transition
P319S phosphorylated	Three state transition	Two-state transition
D321A unphosphorylated	Two-state transition	Two-state transition
D321A phosphorylated	Two-state transition	Two-state transition
D321E unphosphorylated	Two-state transition	Three state transition
D321E unphosphorylated	Three state transition	Two-state transition
D321G unphosphorylated	Three state transition	Two-state transition
D321G phosphorylated	Two-state transition	Two-state transition
D321N unphosphorylated	Two-state transition	Two-state transition
D321N phosphorylated	Three state transition	Two-state transition
D321V unphosphorylated	Two-state transition	Two-state transition
D321V phosphorylated	Three state transition	Two-state transition
E322K unphosphorylated	Two-state transition	Two-state transition
E322K phosphorylated	Three state transition	Three state transition
E322V unphosphorylated	Three state transition	Two-state transition
E322V phosphorylated	Three state transition	Three state transition

In grey the variants that follow the three state transition (Eq.5) in white the variants that follow the two state transition (Eq. 3)

The thermodynamic parameters obtained from the analysis of the far-UV CD and fluorescence changes transitions are reported in Table 7. The GdmCl induced unfolding transitions monitored by far-UV CD are coincident for most of the variants with those obtained from $\overline{\lambda}$ change. Notably, all the variants show conformational stability very similar to that of the wild type.

Table 7 Thermodynamic parameters for GdmCl-induced unfolding equilibrium of ERK2 wild-type and variants measured by far-UV CD and fluorescence spectroscopy.

	$\Delta G^{H}_{2}O$ (kcal/mol)		m (kcal/mol/M)		$[Gdm]_{0.5}$ (M)	
	CD ($[\Theta]_{222}$)	Fluorescence (Lambda Averaged)	CD ($[\Theta]_{222}$)	Fluorescence (Lambda Averaged)	CD ($[\Theta]_{222}$)	Fluorescence (Lambda Averaged)
Wild type unphosphorylated	ΔG_1 = 2.13 ΔG_2 = 5.82	2.79 ± 0.16	m_1 = 1.99 ± 0.30 m_2 = 1.64 ± 0.35	1.34 ± 0.08	$[GdmCl]_{0.5}$ = 1.10 ± 0.04 $[GdmCl]_{0.5}$ = 3.54 ± 0.10	2.08
Wild type phosphorylated	2.30 ± 0.16	2.88 ± 0.20	0.97 ± 0.07	1.19 ± 0.09	2.37	2.42
E33Q unphosphorylated	2.44 ± 0.14	3.33 ± 0.31	0.96 ± 0.06	1.97 ± 0.20	2.54	1.69
E33Q phosphorylated	ΔG_1 = 2.22 ΔG_2 = 9.80	2.98 ± 0.24	m_1 = 1.56 ± 0.01 m_2 = 2.86 ± 0.62	1.30 ± 0.11	$[GdmCl]_{0.5}$ = 1.42 ± 0.03 $[GdmCl]_{0.5}$ = 3.43 ± 0.05	2.29
E81K unphosphorylated	ΔG_1 = 1.10 ΔG_2 = 9.03	ΔG_1 = 2.76 ΔG_2 = 7.02	m_1 = 1.77 ± 0.18 m_2 = 2.98 ± 0.20	m_1 = 2.07 ± 0.24 m_2 = 2.26 ± 0.70	$[GdmCl]_{0.5}$ = 0.62 ± 0.04 $[GdmCl]_{0.5}$ = 3.03 ± 0.08	$[GdmCl]_{0.5}$ = 1.33 ± 0.03 $[GdmCl]_{0.5}$ = 3.10 ± 0.08
E81K phosphorylated	ΔG_1 = 6.04 ΔG_2 = 5.37	ΔG_1 = 5.23 ΔG_2 = 4.72	m_1 = 4.91 ± 0.38 m_2 = 1.49 ± 0.37	m_1 = 4.91 ± 0.37 m_2 = 1.23 ± 0.30	$[GdmCl]_{0.5}$ = 1.23 ± 0.03 $[GdmCl]_{0.5}$ = 3.61 ± 0.14	$[GdmCl]_{0.5}$ = 1.07 ± 0.01 $[GdmCl]_{0.5}$ = 3.84 ± 0.26
L121I unphosphorylated	2.29 ± 0.14	3.27 ± 0.31	1.00 ± 0.07	1.41 ± 0.14	2.29	2.32
L121I phosphorylated	1.73 ± 0.11	2.82 ± 0.21	1.12 ± 0.07	1.46 ± 0.12	1.55	1.93
R135K unphosphorylated	ΔG1 = 2.54 ΔG_2 = 6.47	2.77 ± 0.25	m_1 = 2.28± 0.19 m_2 = 1.96 ± 0.51	1.32 ± 0.13	$[GdmCl]_{0.5}$ = 1.11 ± 0.02 $[GdmCl]_{0.5}$ = 3.30 ± 0.07	2.10
R135K phosphorylated	2.09 ± 0.17	2.38 ± 0.18	0.99 ± 0.07	1.18 ± 0.09	2.12	2.02
D162G unphosphorylated	2.28 ± 0.16	2.67 ± 0.12	1.05 ± 0.08	1.00 ± 0.05	2.17	2.68
D162G phosphorylated	ΔG1 = 3.45 ΔG2 = 2.53	2.11 ± 0.10	m1 = 3.46 ± 0.58 m2 = 1.10 ± 0.12	0.86 ± 0.05	[GdmCl]0.5 = 1.00 ± 0.03 [GdmCl]0.5 = 2.31 ± 0.05	2.45

	ΔG^{H_2O} (kcal/mol)		m (kcal/mol/M)		$[Gdm]_{0.5}$ (M)	
	CD ($[\Theta]_{222}$)	Fluorescence (Lambda Averaged)	CD ($[\Theta]_{222}$)	Fluorescence (Lambda Averaged)	CD ($[\Theta]_{222}$)	Fluorescence (Lambda Averaged)
R191H unphosphorylated	ΔG_1 = 2.65 ΔG_2 = 6.04	5.29 ± 1.06	m_1 = 2.00 ± 0.27 m_2 = 1.56 ± 0.41	2.76 ± 0.55	$[GdmCl]_{0.5}$ = 1.33 ± 0.04 $[GdmCl]_{0.5}$ = 3.87 ± 0.10	1.92
R191H phosphorylated	2.31 ± 0.19	2.55 ± 0.19	1.02 ± 0.09	1.35 ± 0.11	2.27	1.88
L200F unphosphorylated	2.27 ± 0.16	2.88 ± 0.16	0.96 ± 0.07	1.44 ± 0.08	2.36	2.00
L200F phosphorylated	2.15 ± 0.19	3.06 ± 0.30	0.97 ± 0.09	1.72 ± 0.18	2.20	1.77
D235V unphosphorylated	2. 46 ± 0.16	2.48 ± 0.19	1.07 ± 0.08	1.15 ± 0.09	2.30	2.15
D235V phosphorylated	1.77 ± 0.11	2.41 ± 0.23	1.07 ± 0.07	1.10 ± 0.12	1.65	2.18
P319S unphosphorylated	ΔG_1 = 2.23 ΔG_2 = 4.16	ΔG_1 = 3.08 ΔG_2 = 6.44	m_1 = 2.04 ± 0.27 m_2 = 1.17 ± 0.41	m_1 = 2.48 ± 0.23 m_2 = 1.88 ± 0.46	$[GdmCl]_{0.5}$ = 1.09 ± 0.03 $[GdmCl]_{0.5}$ = 3.56 ± 0.20	$[GdmCl]_{0.5}$ = 1.24 ± 0.02 $[GdmCl]_{0.5}$ = 3.43 ± 0.10
P319S phosphorylated	ΔG_1= 3.04 ΔG_2 = 4.30	1.92 ± 0.16	m_1 = 2.79 ± 0.20 m_2 = 1.12 ± 0.34	1.10 ± 0.10	[GdmCl]0.5 = 1.09 ± 0.01 [GdmCl]0.5 = 3.84 ± 0.26	2.02
Y316F unphosphorylated	2.54 ±0.23	2.78 ± 0.25	1.09 ± 0.10	1.39 ± 0.13	2.32	2.00
Y316F phosphorylated	1.46 ± 0.09	2.37 ± 0.19	0.93 ± 0.05	1.09 ± 0.09	1.57	2.17

	ΔG^{H_2O} (kcal/mol)		m (kcal/mol/M)		$[Gdm]_{0.5}$ (M)	
	CD ($[\Theta]_{222}$)	Fluorescence (Lambda Averaged)	CD ($[\Theta]_{222}$)	Fluorescence (Lambda Averaged)	CD ($[\Theta]_{222}$)	Fluorescence (Lambda Averaged)
D321A unphosphorylated	2.16 ± 0.19	2.91 ± 0.20	1.08 ± 0.1	1.39 ± 0.10	2.00	2.09
D321A phosphorylated	1.83 ± 0.14	2.89 ± 0.28	0.90 ± 0.07	1.51 ± 0.15	2.03	1.92
D321E unphosphorylated	2.43 ± 0.20	$\Delta G1$ = 2.93 $\Delta G2$ = 4.76	1.13 ± 0.10	m1 = 2.15 ± 0.15 m2 = 1.41 ± 0.38	2.16	$[GdmCl]_{0.5}$ = 1.36 ± 0.02 $[GdmCl]_{0.5}$ = 3.36 ± 0.13
D321E phosphorylated	ΔG_1 = 3.39 ΔG_2 = 6.99	2.09 ± 0.13	m_1 = 2.50 ± 0.39 m_2 = 1.96 ± 0.57	1.04 ± 0.07	$[GdmCl]_{0.5}$ = 1.59 ± 0.05 $[GdmCl]_{0.5}$ = 3.57 ± 0.08	2.02
D321G unphosphorylated	ΔG_1 = 2.95 ΔG_2 = 4.41	2.91 ± 0.23	m_1 = 2.45 ± 0.43 m_2 = 1.27 ± 0.47	1.37 ± 0.11	$[GdmCl]_{0.5}$ = 1.20 ± 0.04 $[GdmCl]_{0.5}$ = 3.48 ± 0.21	2.12
D321G phosphorylated	2.31 ± 0.13	2.97 ± 0.26	1.01 ± 0.07	1.26 ± 0.12	2.28	2.35
D321N unphosphorylated	2.07 ± 0.16	2.57 ± 0.20	1.00 ± 0.08	1.17 ± 0.10	2.06	2.20
D321N phosphorylated	ΔG_1 = 2.39 ΔG_2 = 8.09	3.23 ± 0.29	m1 = 2.03 ± 0.22 m2 = 2.43 ± 0.71	1.38 ± 0.13	[GdmCl]0.5 = 1.18 ± 0.03 [GdmCl]0.5 = 3.34 ± 0.08	2.35
D321V unphosphorylated	1.26 ± 0.10	2.48 ± 0.18	0.99 ± 0.07	1.08 ± 0.08	1.27	2.29
D321V phosphorylated	ΔG1 = 5.00 ΔG2 = 4.92	2.69 ± 0.22	m1 = 5.02 ± 0.22 m2 = 1.56 ± 0.25	1.32 ± 0.12	[GdmCl]0.5 = 1.00 ± 0.03 [GdmCl]0.5 = 3.16 ± 0.07	2.03
E322K unphosphorylated	1.93 ± 0.14	2.50 ± 0.18	0.76 ± 0.06	1.14 ± 0.09	2.53	2.19
E322K phosphorylated	ΔG_1 = 4.02 ΔG_2 = 3.97	ΔG_1 = 5.08 ΔG_2 = 5.78	m_1 = 4.62 ± 0.49 m_2 = 1.27 ± 0.13	m_1 = 4.73 ± 0.57 m_2 = 1.53 ± 0.52	$[GdmCl]_{0.5}$ = 0.87 ± 0.01 $[GdmCl]_{0.5}$ = 3.13 ± 0.06	$[GdmCl]_{0.5}$ = 1.07 ± 0.02 $[GdmCl]_{0.5}$ = 3.77 ± 0.27
E322V unphosphorylated	ΔG1 = 4.10 ΔG2 = 4.14	2.60 ± 0.19	m1 = 3.63 ± 0.65 m2 = 1.19 ± 0.32	1.12 ± 0.09	[GdmCl]0.5 = 1.13 ± 0.03 [GdmCl]0.5 = 3.48 ± 0.15	2.31
E322V phosphorylated	ΔG_1 = 1.89 ΔG_2 = 5.43	ΔG_1 = 4.99 ΔG_2 = 14.70	m_1 = 2.10 ± 0.28 m_2 = 1.60 ± 0.24	m_1 = 3.67 ± 0.19 m_2 = 4.62 ± 0.15	$[GdmCl]_{0.5}$ = 0.90 ± 0.03 $[GdmCl]_{0.5}$ = 3.39 ± 0.07	$[GdmCl]_{0.5}$ = 1.36 ± 0.01 $[GdmCl]_{0.5}$ = 3.18 ± 0.04

Urea-induced unfolding equilibrium data were measured at 20°C by monitoring ellipticity at 222 nm $[\Theta]_{222}$ and fluorescence intensity averaged emission wavelength $\bar{\lambda}$. ΔG^{H_2O} and *m*

values were obtained from Eq 3-8; $[GdmCl]_{0.5}$ was calculated from Eq 4. Data are reported as the mean ± SE of the fit.

4.5 Enzyme activity assay and kinetic studies

The effect of the mutations on the kinetic properties of ERK2 enzyme was investigated by monitoring the incorporation of phosphate into a peptide substrate (sulfonamido-oxine, Sox AQT0082). This method relays information on the phosphorylation state of proximal serine, threonine or tyrosine residues in peptide- and protein-based kinase substrates. In the absence of phosphorylation, the Sox shows low affinity for Mg (II); upon phosphorylation, Mg (II) affinity is enhanced due to the advantageous chelate effect, involving the Sox and the introduced phosphate group, and fluorescence is turned on. The reaction was performed at 30°C at a fixed concentration of $MgCl_2$.

All the variants interacting with the CD site displayed a reduced catalytic efficiency analysed using AQT0082 over the concentration range 0.069 µM to 40 µM, except for R135K and Y316F that display a four and seventeen-fold increase of activity respectively (Table 8). A decrease of the activity was particularly evident for the variant E81K, D162G and R191H that showed a dramatic decrease of the turnover number (D162G) and an increase of K_M (R191H) in comparison with the values of the wild type. The K_M of the variant R191H was calculated with the double reciprocal method, since the system never reaches the saturation over a substrate concentration interval ranging from 0.69µM to 40µM.

The temperature dependence of the kinase activity of ERK2 wild-type and its variants interacting with the CD site using a low substrate concentration (0.5-1 μM, well below the K_M value) was analysed over the temperature range of 10–45°C (Fig. 13). The optimal temperatures for catalysis were estimated to be 30°C for the wild type, around 35°C for E81K and R135K and around 40°C for D162G and R191H (Table 9 and Fig. 13B). As shown in Table 8, the lowest value of the activation energy ($E_a^{\ddagger}$) is that of the Y316F variant (5.74 kcal·mol^{-1}) which is also the variant displaying a k_{cat} higher than that of the wild-type enzyme (6.38 vs 2.20 s^{-1}) (Table 8). As mentioned before, the activation energies have been measured at a AQT0082 concentration of 0.5-1 μM, well below the K_M value. The activation energy measured in these conditions reflects mainly the enzymatic release of the product and it may be related more to the enzyme flexibility rather than than to its affinity for the substrate.

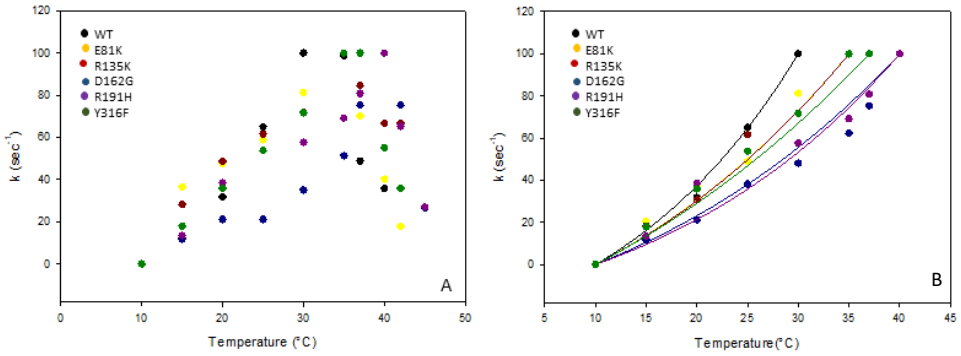

Fig. 13. Effect of temperature on kinase activity of ERK2 wild-type and variants interacting with the CD site. (A) Temperature dependence of kinase activity of ERK2 wild-type and variants. (B) Non-linear fit of the temperature dependence of ERK2 activity to the Arrhenius equation: $k = Ae\text{-}Ea_{\ddagger}^{\ddagger}/RT$ (1). Assays were performed under the conditions described in Materials and Methods.

Table 8 Kinetic parameters of ERK2 wild-type and variants interacting with CD site.

	K_M (μM)	[peptide]	k_{cat} (s^{-1})	[enzima]	k_{cat} / K_M (s^{-1} μM^{-1})	Vmax (μM x s^{-1})	Specific activity (μM x s^{-1} μg^{-1})
Wild type	1.81 ± 0.29	0.5 μM	2.20	5.0 nM	1.215	$1.101*10^{-2}$	1.218
E81K	2.64 ± 0.72	0.5 μM	1.24	10.0 nM	0.470	$1.241*10^{-2}$	0.748
R135K	6.19 ± 1.90	0.5 μM	28.34	0.5 nM	4.578	$1.417*10^{-2}$	17.082
D162G	1.43 ± 0.47	0.5 μM	$9.30*10^{-4}$	1437 nM	$6.503*10^{-4}$	$1.343*10^{-3}$	0.001
R191H	50.0 ± 2.07	1.0 μM	$2.00*10^{-2}$	2381 nM	$3.636*10^{-4}$	$4.835*10^{-2}$	0.012
Y316F	1.57 ± 0.30	0.5 μM	6.83	1.0 nM	4.350	$6.829*10^{-3}$	4.116

ERK2 activity was determined at 30°C, with AQT0082 and 5 mM MgATP, as described in Materials and Methods. Kinetic parameters for ERK2 activity were determined at 30°C by using at least at 10 different concentrations of AQT0082**.** Data are reported as the mean ± SE of the fit.

Table 9 Effect of temperature on kinase activity of ERK2 wild-type and variants interacting with CD site.

	T_{max} (°C)	*E*a ‡ (kcal/mol)
Wild type	30.0	9.95 ± 0.41
E81K	35.0	6.90 ± 0.95
R135K	35.0	7.02 ± 0.90
D162G	40.0	6.94 ±0.96
R191H	40.0	8.09 ± 0.99

Y316F	37.0	5.74 ± 0.26

$E_a^{a‡}$ was determined by the Arrhenius equation ($k = A$e- $Ea‡/RT$) in the temperature range between 10 °C and the optimal temperature of each protein.

Most of the variants in the CD site displayed a reduced catalytic efficiency analysed using AQT0082 over the concentration range 0.069 µM to 40 µM, with the exception of D321E and E322V that display a nineteen and two fold increase of activity respectively (Table 10) (Fig. 14A). The variant D321A shows a specific activity of 3.571 µM x s^{-1} μg^{-1}. The decrease of the activity was particularly evident for the variant D321G and E322K, as revealed by the high concentration of the enzyme and the K_M increase.
The variant D321N reveals a lower catalytic efficiency, with small differences for the value K_M and k_{cat} compared with the wild type.

The variants in the CD site display an optimal temperature for catalysis lower than wild type, except for the variant D321V with a temperature around 25°C (Fig. 14B) (Table 11). Notably (Table 10), the variants with the highest specific activity (D321E and E322V) shows the lower value of the $E_a^‡$ (5.24 kcal·mol^{-1} and 3.99 kcal·mol^{-1}) (Table 11).

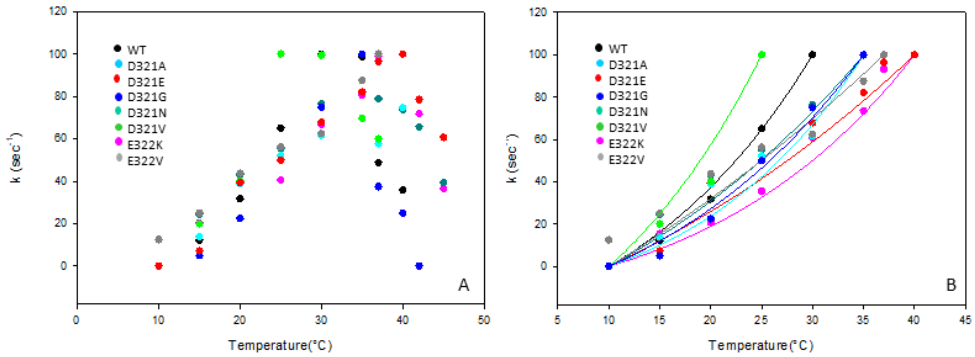

Fig. 14. Effect of temperature on kinase activity of ERK2 wild-type and variants in the CD site. (A) Temperature dependence of kinase activity of ERK2 wild-type and variants. (B) Non-linear fit of the temperature dependence of ERK2 activity to the Arrhenius equation Equation: $k = Ae^{-Ea‡/RT}$ (1). Assays were performed under the conditions described in Materials and Methods.

Table 10 Kinetic parameters of ERK2 wild-type and variants in the CD site.

	K_M (μM)	[peptide]	k_{cat} (s^{-1})	[enzima]	k_{cat} / K_M (s^{-1} μM^{-1})	Vmax (μM x s^{-1})	Specific activity (μM x s^{-1} μg^{-1})
Wild type	1.81 ± 0.29	0.5 μM	2.20	5.0 nM	1.215	$1.101*10^{-2}$	1.218
D321A	7.42 ± 2.69	0.5 μM	5.93	10 nM	0.799	$5.925*10^{-2}$	3.571
D321E	7.11 ±1.77	0.5 μM	32.96	0.5 nM	4.635	$1.648*10^{-2}$	19.867
D321G	0.65 ± 0.12	0.1 μM	$1.64*10^{-4}$	2000 nM	$2.523*10^{-3}$	$3.278*10^{-3}$	0.001
D321N	2.95 ± 0.76	0.5 μM	1.26	5.0 nM	0.427	$6.307*10^{-3}$	0.760
D321V	3.47 ± 0.55	0.5 μM	2.41	5.0 nM	0.694	$1.207*10^{-2}$	1.455
E322K	4.70 ± 0.81	0.5 μM	1.64	5.0 nM	0.349	$8.223*10^{-3}$	0.099
E322V	2.81 ± 0.68	0.5 μM	3.80	5.0 nM	1.352	$1.903*10^{-2}$	2.294

ERK2 activity was determined at 30°C, with AQT0082 and 5 mM MgATP, as described in Materials and Methods. Kinetic parameters for ERK2 activity were determined at 30°C by using at least at 10 different concentrations of AQT0082. Data are reported as the mean ± SE of the fit

Table 11 Effect of temperature on kinase activity of ERK2 wild-type and variants in the CD site.

	T_{max} (°C)	*Ea* ‡ (kcal/mol)
Wild type	30.0	9.95 ± 0.41
D321A	35.0	11.43 ± 1.40
D321E	40.0	5.24 ± 0.39
D321G	35.0	8.97 ± 0.50
D321N	35.0	6.67 ± 0.49
D321V	25.0	10.12 ± 1.49
E322K	40.0	9.76 ± 1.58
E322V	37.0	3.99 ± 0.26

$E_a^{a‡}$ was determined by the Arrhenius equation ($k = Ae^{-Ea‡/RT}$) in the temperature range between 10 °C and the optimal temperature of each protein.

The variants in others site of the protein show a reduced activity compared to that of the wild type, with the exception of E33Q, whose the observed specific activity is comparable to that of the wild type (Table 12) (Fig. 15A). The variants L121I and P319S display an optimal temperature at 35°C and the variant D235V at 37°C (Table 13). The variants, E33Q and L200F show an optimal temperature for catalysis of 25°C and 30°C respectively(Table 13).

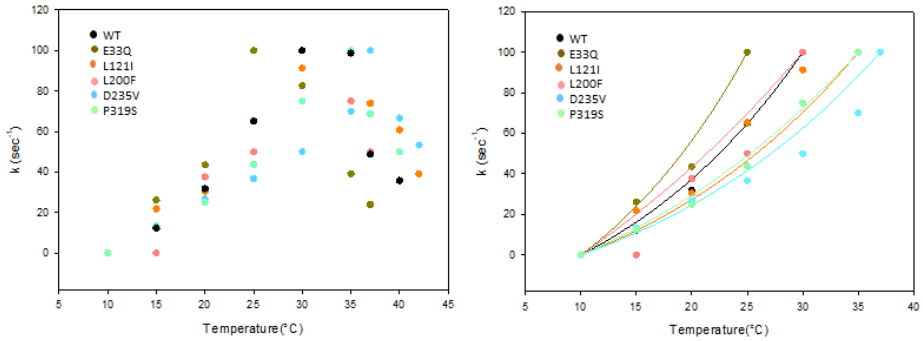

Fig. 15. Effect of temperature on kinase activity of ERK2 wild-type and variants in other site of the protein. (A) Temperature dependence of kinase activity of ERK2 wild-type and variants. (B) Non-linear fit of the temperature dependence of ERK2 activity to the Arrhenius equation Equation: $k = Ae^{-Ea^{\ddagger}/RT}$ (1). Assays were performed under the conditions described in Materials and Methods.

Table 12 Kinetic parameters of ERK2 wild-type and variants in other site of the protein.

	K_M (µM)	[peptide]	k_{cat} (s^{-1})	[enzima]	k_{cat} / K_M (s^{-1} μM^{-1})	Vmax (µM x s^{-1})	Specific activity (µM x s^{-1} μg^{-1})
Wild type	1.81 ± 0.29	0.5 µM	2.20	5.0 nM	1.215	$1.101*10^{-2}$	1,218
E33Q	3.90 ± 0.67	0.5 µM	1.98	5.0 nM	0.508	$9.913*10^{-3}$	1,195
L121I	7.12 ± 2.09	0.5 µM	$7.20*10^{-2}$	390.0 nM	0.010	$2.798*10^{-2}$	0,043
L200F	3.23 ± 0.40	0.5 µM	$2.43*10^{-2}$	500 nM	$7.523*10^{-3}$	$1.215*10^{-2}$	0,015
D235V	3.97 ± 0.67	0.5 µM	$3.10*10^{-1}$	50 nM	0.078	$1.550*10^{-2}$	0,187
P319S	4.26 ±1.11	0.5 µM	$3.62*10^{-1}$	100 nM	0.085	$3.622*10^{-2}$	0,218

ERK2 activity was determined at 30°C, with AQT0082 and 5 mM MgATP, as described in Materials and Methods. Kinetic parameters for ERK2 activity were determined at 30°C by using at least at 10 different concentrations of AQT0082. Data are reported as the mean ± SE of the fit

Table 13 Effect of temperature on kinase activity of ERK2 wild-type and variants in others site of the protein.

	Tmax (°C)	*E*a ‡ (kcal/mol)
Wild type	30.0	9.95 ± 0.41
E33Q	25.0	11.11± 1.24
L121I	35.0	8.97 ± 1.4
L200F	30.0	5.76 ± 0.86
D235V	37.0	8.60 ± 0.88
P319S	35.0	7.56 ±0.31

$E_a^{a‡}$ was determined by the Arrhenius equation $k = Ae^{-Ea‡/RT}$ in the temperature range between 10 °C and the optimal temperature of each protein.

4.6 DSF analysis: the thermal stabilization of the native protein structure upon ligand binding

The differential scanning fluorimetry(DSF) is an assay that allows to monitor the denaturation of the protein by measuring the corresponding changes in fluorescence emission upon temperature increase.

A change in the protein environment and/or complex formation with other molecules (ATP, drugs, inhibitors, ligand…) can stabilize a protein through a reduction of the Gibbs free energy of the complex, resulting from the creation of new molecular interactions (hydrogen bonds, van der Waals interactions, etc.) or conformational reordering of the target protein. This decrease in the Gibbs free energy results in an increase in thermal stability and thereby an increase in the melting temperature (Tm) [Gao K. et al, 2020].

This analysis was carried out trying to answer the specific question, if the mutation in the variants might interfere with the binding with known inhibitors. All the inhibitors occupy part of the adenine binding pocket and form hydrogen bonds with the hinge region connecting the small and large lobes of the enzyme (Table 14).

Table 14 Table of the selected inhibitors of ERK2. The table show all the inhibitors analysed by DSF for ERK2 and its variants.

Compounds	Supplier	Compound Description	References
SCH772984	Cayman	ERK inhibitor	Morris E.J. et al, 2013
MK-8353	Selleckchem	ERK inhibitor	Boga S.B. et al, 2018
AZD0364	Cayman	ERK inhibitor	Simpson I. et al, 2018
KO-947	Cayman	ERK inhibitor	Burrows S. et al, 2017
LY3214996	Cayman	ERK inhibitor	Bhagwat S. V. et al, 2020
FR 180204	Cayman	ERK inhibitor	Sreekanth G. P. et al, 2014
GDC-0994	Cayman	ERK inhibitor	Varga A. et al, 2019
Ulixertinib	Cayman	ERK inhibitor	Sullivan R. J. et al, 2018
Pexmetinib	Cayman	A dual inhibitor of Tie2 and p38 MAPK	Chiavarina B. et al, 2021
Magnolin	Cayman	ERK inhibitor	Xu K. et al, 2020
Vx-11e	Cayman	ERK inhibitor	Jasek-Gajda E. et al,2019

All the variants show T_m values similar to that of the wild type with a small difference of 2-4 °C, (Table 15). The ligand binding stabilizes the native state resulting in an increase of the T_m value. All the variants show small

divergences compared to the wild type (Table 2) which could be associated with differences in the thermal stability between wild type and variants, as we reported in the paragraph 4.3. The values below 1 suggest that the ligand binds to the protein preferentially to the unfolded state or non-native states and stabilizes (increases the molar fraction or the population of) the unfolded or non-native states.

It was impossible to analyse the data for some inhibitors (Vx-11e, Pexmetinib and MK-8353), since the changes in fluorescence emission is due to the intrinsic characteristics of ligands that are considered fluorophores themselves and could interfere with analyses.

Table 15 Effect of the binding with inhibitors on melting temperature of ERK2 wild-type and variants.

A

	ΔTm [°C] WT	ΔTm [°C] E81K	ΔTm [°C] R135K	ΔTm [°C] D162G	ΔTm [°C] R191H	ΔTm [°C] Y316F
AZD0364	12,8 ± 0,05	13±0,11	13,3±0,34	11,1±1,18	11,6±0,24	11,9±0,57
KO-947	11,8 ± 0,22	12,2±0,06	12,8±0,48	10±1,16	10,7±0,21	11,1±0,59
SCH772984	11,6 ± 0,27	11,8±0,16	11,8±0,23	9,7±1,47	10,2±0,07	10,1±0,58
GDC-0994	10,3 ± 0,3	11,2±0,4	11±0,61	7,8±1,65	9,4±0,27	9,4±0,37
LY3214996	9,4 ± 0,21	9,9±0,12	10,2±0,54	7,5±1,24	8,6±0,12	8,8±0,23
Ulixertinib	7,9 ± 0,39	8,4±0,21	8,2±0,2	5,6±1,42	6,8±0,12	6,8±0,26
FR 180204	2,8 ± 1,27	3,6±0,42	3,6±0,43	0,4±2,44	2,7±0,72	1,7±0,27
Magnolin	0,2 ± 0,11	0,2±0,17	0,4±0,4	-3,4±2,48	-0,5±0,31	-1,6±0,6

B

	ΔTm [°C] WT	ΔTm [°C] D321A	ΔTm [°C] D321E	ΔTm [°C] D321G	ΔTm [°C] D321N	ΔTm [°C] D321V	ΔTm [°C] E322K	ΔTm [°C] E322V
AZD0364	12,8 ± 0,05	13,3 ± 0,51	13,7 ± 1,03	13,9±0,89	14,5 ± 0,21	13±0,03	10,2 ± 0,01	13,6±0,7
KO-947	11,8 ± 0,22	12,3 ± 0,37	12,8 ± 1,01	13,1±0,86	13,6 ± 0,23	11,8±0,02	9,2 ± 0,08	12,4±0,35
SCH772984	11,6 ± 0,27	11 ± 0,09	11,7 ± 1,15	12,2±0,89	12,7 ± 0,24	11,6±0,13	8,7 ± 0,15	12,2±0,66
GDC-0994	10,3 ± 0,3	10,7 ± 0,2	11,2 ± 0,74	11,8±0,58	12,2 ± 0,15	10,3±0,28	6,8 ± 0,49	11,2±0,5
LY3214996	9,4 ± 0,21	9,8 ± 0,02	10,3 ± 0,88	10,8±0,67	11,2 ± 0,14	10±0,02	6,6 ± 0,01	10,3±0,49
Ulixertinib	7,9 ± 0,39	8,3 ± 0,31	8,8 ± 0,95	9,1±0,79	9,5 ± 0,24	8,3±0,09	4,6 ± 0,14	8,5±0,16
FR 180204	2,8 ± 1,27	3,2 ± 0,2	3,7 ± 0,86	4,6±0,94	5,4 ± 0,58	5±0,44	-1,3 ± 0,08	3,4±1,25
Magnolin	0,2 ± 0,11	0,6 ± 0,24	1,4 ± 1,43	1,9±1,13	2,6 ± 0,35	0,3±0,1	-5,3 ± 0,35	0,6±0,24

C

	ΔTm [°C] WT	ΔTm [°C] E33Q	ΔTm [°C] L121I	ΔTm [°C] L200F	ΔTm [°C] D235V	ΔTm [°C] P319S
AZD0364	12,8 ± 0,05	11,7±0,31	11,6±0,24	11,7±0,36	13,4±0,63	14,1±0,89
KO-947	11,8 ± 0,22	11,4±0,47	10,7±0,22	10,8±0,14	11,5±0,94	13,8±0,69
SCH772984	11,6 ± 0,27	11,2±0,44	10,3±0,07	10,3±0,14	11,4±0,31	12,5±0,81
GDC-0994	10,3 ± 0,3	9,9±0,34	9,3±0,25	9,2±0,22	9,6±0,37	10,7±0,55
LY3214996	9,4 ± 0,21	8,6±0,17	8,6±0,12	8,5±0,09	9,1±0,21	10,2±0,47
Ulixertinib	7,9 ± 0,39	7,1±0,06	6,9±0,13	6,9±0,25	7,3±0,3	8,5±0,51
FR 180204	2,8 ± 1,27	2,7±0,27	3,2±0,16	2,1±0,4	2,6±0,16	3,5±0,09
Magnolin	0,2 ± 0,11	-0,4±0,05	-0,2±0,13	-0,9±0,25	0,2±0,61	0,6±0,35

The Buffer used to dilute the protein was 10 mM HEPES pH 7.5 and 500 mM NaCl The proteins (2µM) were mixed with 12 µM inhibitors and SyPRO Orange at 1:1000 dilution. (A) DSF analysis of the variants interacting with CD site. (B) DSF analysis of the variants in the CD site. (C) DSF analysis of the variants in other site of the protein.

4.7 Isothermal titration calorimetry (ITC) study on the binding of inhibitors GDC-0994 to the variant E322K

The most studied ERK2 variant is E322K, identified in squamous cell carcinomas and identified as a cancer mutation [Taylor IV C. et al, 2019]. Little is known about the cellular consequences associated with this mutation, other than an apparent increase in tumour resistance to inhibitors. Isothermal titration calorimetry (ITC) assay was performed to quantify the effect of this mutation on the binding with the inhibitor. ITC is a convenient and widely used approach to directly measure the amount of heat released or absorbed during association processes of biomolecules in solution and to quantitatively estimate the interaction affinity between proein and its ligands.

The profile of the ERK2 wild type titrated with the inhibitor GDC-0994 shows an affinity value at about 3.8 nM (Fig.16A), in the same order of

magnitude with that of the variant E322K at 7.9 nM (Fig. 16B). Besides confirming direct interactions with the target of interest, thermodynamic measurements also provide insight into the nature of the noncovalent forces responsible for binding. Polar interactions tend to contribute favourably to the enthalpic component, whereas entropically favoured interactions tend to be more hydrophobic. The binding of ERK2 wild type and the inhibitor GDC-0994 show an enthalpy-driven interaction, driven by hydrogen-bonding and van der Waals forces. On the other hand, interaction of ERK2 E322K and the inhibitor GDC-0994 display an entropy-driven interaction, mainly characterized by hydrophobic forces.

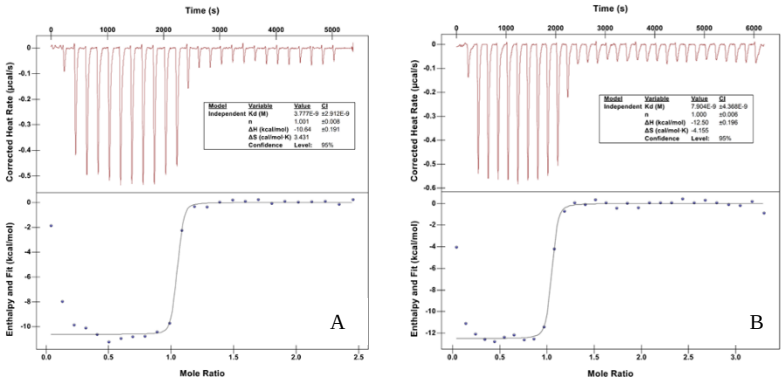

Fig. 16. Effect of the binding with GDC-0994 on the amount of heat released or absorbed during association of ERK2 wild-type (A) and ERK2 E322K (B). The measurements were performed using an VP-ITC (MicroCal) with 20µM of inhibitor in the cell and 170 µM of ERK2 wild type in the syringe at 30 °C.

4.8 An isothermal titration calorimetry study on the binding of DUSP6 and MEK2 peptides to the D321N

We performed an ITC assay to understand the role of the mutation D321N. This variant shows an increased activity in cells and evades inactivation by dual-specificity phosphatases [Taylor C. A. et al, 2019] such as DUSP6. The DUSP6 plays a crucial tumour-suppressive role via pivotal negative-feedback regulation of the ERK2 in the MAPK signalling pathway [Zhang Z. et al, 2010; Furukawa T. et al, 2003; Chan D.W. et al, 2008]. The Kinase interaction motif (KIM) GIMLRRLQKGNLPVRAL of DUSP6 directly binds to the docking site with high-affinity association. The ITC analysis between the P-ERK2 wild type and the peptide (KIM) DUSP6 show a K_D of 14.4 µM (Fig. 17). This analysis shows an enthalpy-driven interaction hydrogen-bonding and van der Waals interactions.

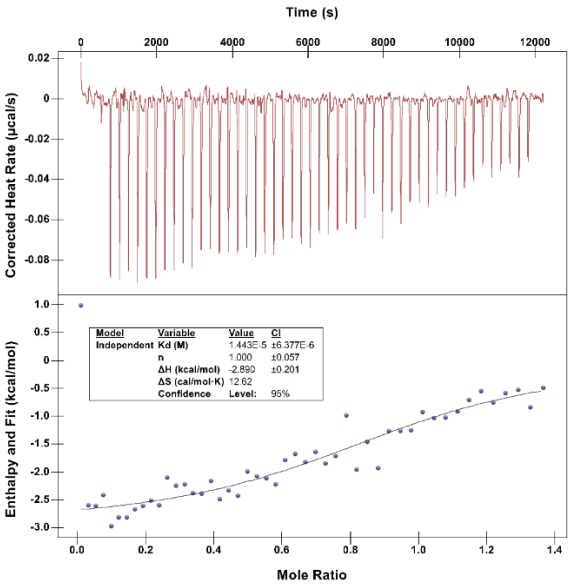

Fig. 17. Effect of the binding with pepDUSP6 on the amount of heat released or absorbed during association of P-ERK2 wild-type. The measurements were performed using an "Affinity ITC" (TA-Instruments) with 500µM of peptide (pepDUSP6) in the syringe and 93 µM of P- ERK2 wild type in the cell at 30 °C.

The analysis of P-ERK2 variant D321N (Fig. 18) shows a no detectable binding heat suggesting that there is no binding with the peptide (pepDUSP6), because the negative charge of the aspartate, required for docking, is lost in the asparagine variant.

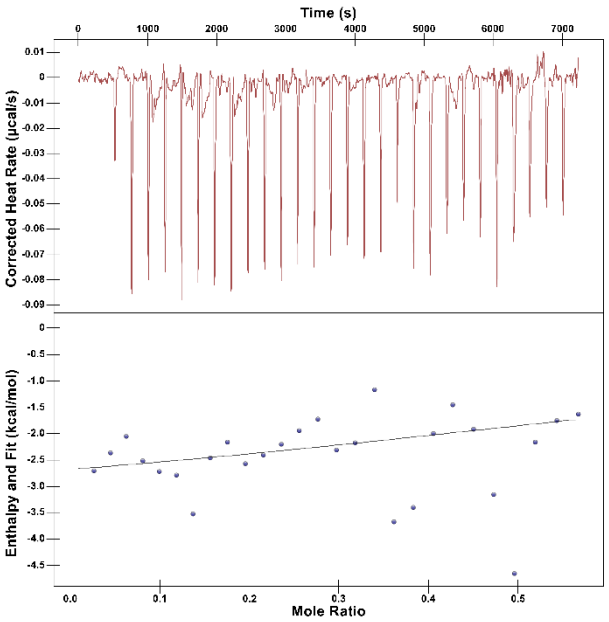

Fig. 18. Effect of the binding with pepDUSP6 on the amount of heat released or absorbed during association of P-ERK2 D321N variant. The measurements were performed using an "Affinity ITC" (TA-Instruments) with 500µM of peptide(pepDUSP6) in the syringe and 113 µM of P-ERK2 D321N in the cell at 30 °C.

We also analysed the binding with the pepMEK2 of ERK2 wild type compared with that to the variant D321N. PepMEK2 is a MAP Kinase ERK2 substrate docking peptide derived from MLARRKPVLPALTINP. The PepMEK2 sequence is located remotely from the active site of the substrate or processing enzyme and causes conformational changes in ERK2 that make phosphorylation sites more available. The ITC profiles show the interaction of NP-ERK2 wild type with pepMEK2 (Fig. 19) and this titration show an enthalpy-driven interaction with a K_D value of 6.1 µM.

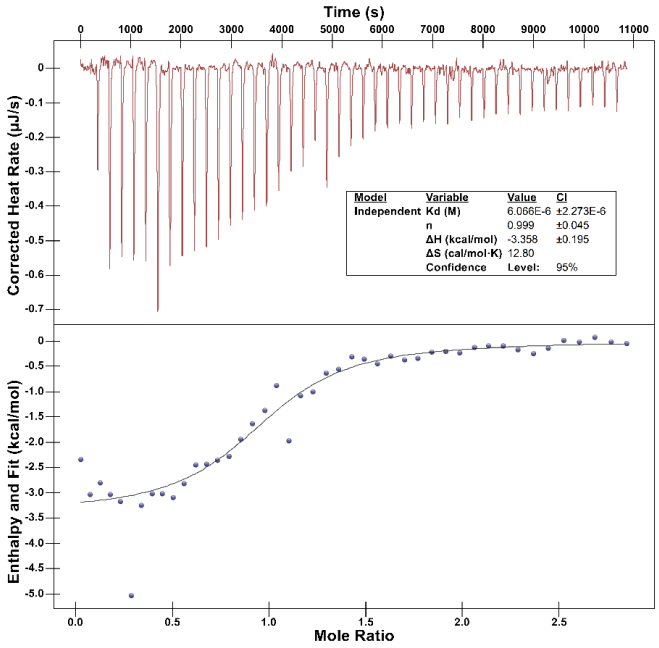

Fig. 19. Effect of the binding with pepMEK2 on the amount of heat released or absorbed during association of NP-ERK2 wild-type. The measurements were performed using an "Affinity ITC" (TA-Instruments) with 500µM of peptide (pepMEK2) in the syringe and 145 µM of NP-ERK2 wild type in the unphosphorylated form in the cell at 30 °C.

The titration profile of the pepMEK2 in the protein NP-ERK2 D321N (K_D of 160 µM) (Fig. 20) display an entropy-driven interaction more hydrophobic in character compared to enthalpy-driven interaction. These results confirm that in D321N structural and charge change in the CD-site can interfere the interaction with the substrate.

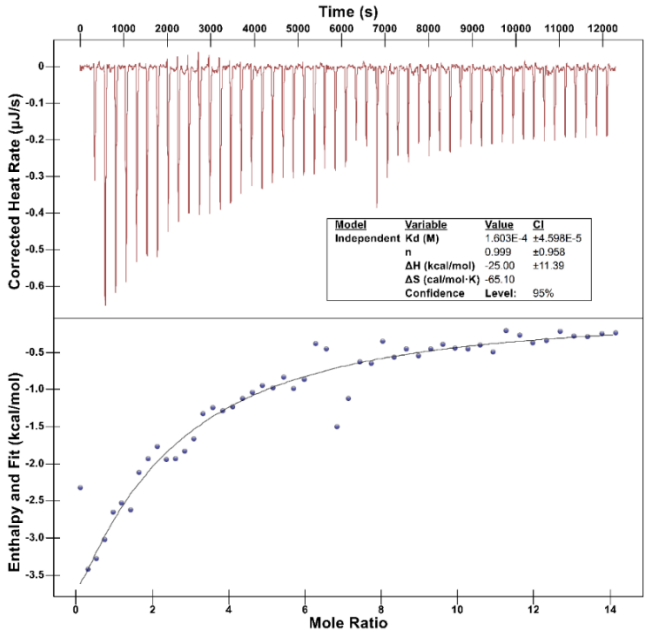

Fig. 20. Effect of the binding with pepMEK2 on the amount of heat released or absorbed during association of NP-ERK2 D321N. The measurements were performed using an "Affinity ITC" (TA-Instruments) with 500 µM of peptide (pepMEK2) in the syringe and 145 µM of NP-ERK2 D321N variant in the cell at 30 °C.

4.9 Crystallization of ERK2 D321N in its phosphorylated form

The rational developmente of inhibitor requires structural information of ERK2 D321N in its active form. Thus, we worked extensively on crystallization techniques, including seeding and soaking, to improve the quality of the crystals for data collection.

Fig. 21. Crystals of P-ERK2 D321N. The protein ERK2 D321N was concentrated to 9 mg/mL and incubated for 1h with 1 mM of AZD0364 (Cayman Item No. 29827) and crystallized using the reservoir solutions containing 1.6M MgSO4 and 0.1 M MES *p*H 6.0. D321N crystals were small needle clusters and after 1 week harvested.

Fig. 22. The human P-ERK2 D321N.

The structure analyzed by Dr Apirat Chaikuad at the SGC in Frankfurt (Goethe University) gave a resolution of 1.8 Å. As shown in Figure 22 in our structure the activation loop seems partially disordered with no clear density for the phosphorylated TXY.

The A-loop compared in all structures has very different conformation (Fig. 23) in our structure, it is possible that the inhibitor causes some local rearrangements in the insertion site.

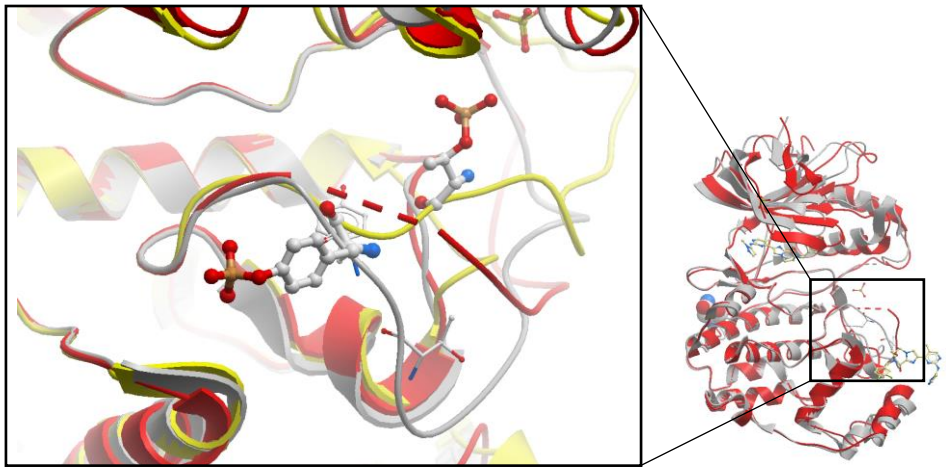

Fig. 23. Comparison among three different protein one in red is the human P-ERK2 D321N, in yellow the rat P-ERK2 wild type(PDB: 2ERK) and in grey the rat NP-ERK2 D321N (PDB: 6OT6).

Discussion

Missense mutations that generate protein variants with a single amino acid variation (SAV), are of particular interest in biomedicine, since a single amino acid substitution may induce drastic structural alterations, which compromise the protein stability, or may induce crucial structural alterations able to perturb binding interfaces or/ and the protein function. For most of the missense mutations, their impact at structural and function level of the wild type protein is still unknown.

In this thesis a biochemical characterization of SAV that have been detected in different types of cancer has been performed to investigate the effect of missense substitutions on the nsSNV structure, function and stability and to contribute to their phenotypic characterization.

Our study focuses on ERK2 single amino acid somatic variants that participate in the Ras-Raf-MEK-ERK signal transduction cascade. This cascade is involved in the regulation of a large variety of processes including cell adhesion, cell cycle progression, cell migration, cell survival, differentiation, metabolism, proliferation, and transcription.

The data reported here provide an important contribution to the collection of structural, function and stability data on disease-related protein variants. Our analysis of a dataset of 17 nsSNVs NP-ERK2 and P-ERK2 variants indicates that the single amino acid substitutions can affect the properties of the native protein in different ways (Table1). Most of the NP-ERK2 and P-ERK2 variants studied show significantly alterations of their tertiary arrangements (Fig. 3-5A, B), as indicated by the differences in their NEAR-UV CD fine structure and/or in the ellipticity amplitude, when compared to the corresponding wild-type protein (Fig. 2A). The changes in tertiary arrangements of most of the variants were also evident from their intrinsic fluorescence spectra (Fig.3-5C, D). Notably, the mutations do not cause large conformational changes in the protein 3D structure, but only local changes and/or tertiary alterations involving the mutated residues and their chemical surrounding (Fig. 2-6). Due to the usually surface exposed location of the mutated residues (in ERK2 71% are solvent exposed, 29% are buried) (Table 1), the consequences of the missense mutations were not obvious, and the most important finding is that these local changes may be responsible for local differences in protein flexibility that may affect protein-protein interactions.

Hence, a local effect may become global when is reflected on the overall structural and functional properties of the protein, even though they are restricted to a narrow protein environment, without a significant perturbation of the protein 3D structure.

The global effects of local alterations, "GLOCAL", are evident from the analysis of the thermal unfolding process. The evaluation of the T_m values has shown some differences compared with those of the wild type (Table 3-5). Some variants in the CD-site, such as D321N and E322K, show huge differences in thermal stability, due not only to the point mutations, but also to the position of mutations within the structure.

However, the D321N mutation is important for substrate binding and its thermodynamic instability could lead to the loss of the protein biological function. In addition, post-translational modifications e.g., phosphorylation, in the variants and in the wild type may produce different effects.

Independently from the position of the mutation on the protein, one of the results of the amino acid substitutions, observed in our study is a significant change in the catalytic activity of the variants and most of the variants display an altered catalytic efficiency (Table 8-13). The variants can be divided in those with a reduced and those with higher catalytic efficiency, in comparison with the wild type and they may drive differential responses to MAPK-pathway-directed therapies, providing a reference for future treatment decisions. The ERK2 variant D321N has been extensively characterized and it has been reported that causes increased activity in cells and evades inactivation by dual-specificity phosphatases. In our study, using Chelation-Enhanced Fluorescence, the specific activity of P-ERK2 D321N is lower than the wild type. The reduced catalytic properties of other ERK2 variants (P-ERK2 D321G, P-ERK2 E322K) may suggest that their expression in the

neoplastic cells would significantly burden the metabolic flux through the RAS-RAF-MEK-ERK pathway.

The possibility of ERK2 D321N to keep the pathways RAS-RAF-MEK-ERK always in an active state was confirmed by ITC assay using the peptide DUSP6, which does not interact with the variant. If ERK2 D321N variant is refractory to negative regulation by DUSP6, this might lead to sustain the hyperactivation of ERK2 signal (Fig. 17-18).

The structure of the variant ERK2 D321N was determined by X-ray crystallography and showed significant changes that may affect its functional properties. In the structure, the activation loop seems partially disordered with no clear density for the phosphorylated TXY (Fig. 22). The A-loop compared in all structures have very different conformation (Fig. 23), in our case a possible reason could be that the inhibitor co-crystallized causes some local rearrangements in the insertion site.

In conclusion our results indicate that the effect of the mutation observed in cancer tissues are directed to local changes of tertiary structure. More importantly, they do not affect binding to type I kinase inhibitors studied. This study can be helpful to reduce the informational gap in the analysis of nsSNVs that are found in different cancer tissues, especially in the search of new small molecules able to selectively interact with the variants, which is an essential preliminary step to personalized medicine.

Chapter 6

Attachments

HHS Public Access
Author manuscript
Hum Mutat. Author manuscript; available in PMC 2020 September 01.

Published in final edited form as:
Hum Mutat. 2019 September ; 40(9): 1400–1413. doi:10.1002/humu.23789.

Characterization of human frataxin missense variants in cancer tissues

Maria Petrosino[1,2,3], **Alessandra Pasquo**[4], **Leonore Novak**[1], **Angelo Toto**[1,5], **Stefano Gianni**[1,5], **Elide Mantuano**[6], **Liana Veneziano**[6], **Velia Minicozzi**[7], **Annalisa Pastore**[8], **Rita Puglisi**[8], **Emidio Capriotti**[9], **Roberta Chiaraluce**[*,1], **Valerio Consalvi**[1]

[1]Dipartimento di Scienze Biochimiche "A. Rossi Fanelli". Sapienza University of Rome, Rome, Italy

[2]Current address: IRCCS Istituto Neurologico Carlo Besta, Milano, Italia

[3]European Brain Research Institute-Fondazione Rita Levi Montalcini, Roma, Italia.

[4]ENEA CR Frascati, Diagnostics and Metrology Laboratory,FSN-TECFIS-DIM, Frascati, Italy

[5]Istituto di Biologia e Patologia Molecolari del CNR, Sapienza Università di Roma, Rome, Italy

[6]Institute of Translational Pharmacology, CNR, Rome, Italy

[7]INFN and Department of Physics, University of Rome Tor Vergata, Rome, Italy

[8]The Wohl Institute, King's College London, London, United Kingdom

[9]Department of Pharmacy and Biotechnology (FaBiT), University of Bologna, Bologna, Italy

Abstract

Human frataxin is an iron binding protein involved in the mitochondrial Fe-S clusters assembly, a process fundamental for the functional activity of mitochondrial proteins. Decreased level of frataxin expression is associated with the neurodegenerative disease Friedreich ataxia. Defective function of frataxin may cause defects in mitochondria, leading to increased tumorigenesis. Tumour initiating cells show higher iron uptake, a decrease in iron storage and a reduced Fe-S clusters synthesis and utilization. In this study we selected, from COSMIC database, the somatic human frataxin missense variants found in cancer tissues p.D104G, p.A107V, p.F109L, p.Y123S, p.S161I, p.W173C, p.S181F, and p.S202F to analyze the effect of the single amino acid substitutions on frataxin structure, function and stability. The spectral properties, the thermodynamic and the kinetic stability, as well as the molecular dynamics of the frataxin missense variants found in cancer tissues point to local changes confined to the environment of the mutated residues. The global fold of the variants is not altered by the amino acid substitutions, however some of the variants show a decreased stability and a decreased functional activity in comparison to that of the wild type protein.

[*]Corresponding author: roberta.chiaraluce@uniroma1.it.

CONFLICT OF INTERESTS

The authors declare that there are no conflict of interests.

International Journal of
Molecular Sciences

MDPI

Review

Analysis and Interpretation of the Impact of Missense Variants in Cancer

Maria Petrosino [1], Leonore Novak [1], Alessandra Pasquo [2], Roberta Chiaraluce [1], Paola Turina [3], Emidio Capriotti [3,*] and Valerio Consalvi [1,*]

[1] Dipartimento Scienze Biochimiche "A. Rossi Fanelli", Sapienza University of Rome, 00185 Roma, Italy; maria.petrosino@uniroma1.it (M.P.); leonore.novak@uniroma1.it (L.N.); roberta.chiaraluce@uniroma1.it (R.C.)
[2] ENEA CR Frascati, Diagnostics and Metrology Laboratory FSN-TECFIS-DIM, 00044 Frascati, Italy; alessandra.pasquo@enea.it
[3] Dipartimento di Farmacia e Biotecnologie (FaBiT), University of Bologna, 40126 Bologna, Italy; paola.turina@unibo.it
* Correspondence: emidio.capriotti@unibo.it (E.C.); valerio.consalvi@uniroma1.it (V.C.)

Abstract: Large scale genome sequencing allowed the identification of a massive number of genetic variations, whose impact on human health is still unknown. In this review we analyze, by an in silico-based strategy, the impact of missense variants on cancer-related genes, whose effect on protein stability and function was experimentally determined. We collected a set of 164 variants from 11 proteins to analyze the impact of missense mutations at structural and functional levels, and to assess the performance of state-of-the-art methods (FoldX and Meta-SNP) for predicting protein stability change and pathogenicity. The result of our analysis shows that a combination of experimental data on protein stability and in silico pathogenicity predictions allowed the identification of a subset of variants with a high probability of having a deleterious phenotypic effect, as confirmed by the significant enrichment of the subset in variants annotated in the COSMIC database as putative cancer-driving variants. Our analysis suggests that the integration of experimental and computational approaches may contribute to evaluate the risk for complex disorders and develop more effective treatment strategies.

Keywords: protein structure; protein stability; protein function; single amino acid variant; putative cancer driving variant; free-energy change

Citation: Petrosino, M.; Novak, L.; Pasquo, A.; Chiaraluce, R.; Turina, P.; Capriotti, E.; Consalvi, V. Analysis and Interpretation of the Impact of Missense Variants in Cancer. *Int. J. Mol. Sci.* **2021**, *22*, 5416. https://doi.org/10.3390/ijms22115416

Academic Editors: Giovanni Minervini and Emanuela Leonardi

Received: 30 March 2021
Accepted: 17 May 2021
Published: 21 May 2021

1. Introduction

Recent advances in the high-throughput sequencing technologies have led to the detection of a large amount of genomic data, which are used for generating detailed catalogues of genetic variations in both diseased and healthy patients [1,2]. Such genetic differences are at the basis of distinctive traits associated with the susceptibility to specific disease and/or drug response [3]. The most common genetic differences in the human genome are single nucleotide polymorphisms (SNPs), which are defined as single nucleotide variations (SNVs) occurring with a frequency of more than 1% in the population [4,5]. These differences occur on average once every 300–400 base pairs [6], either in coding or in non-coding regions (Figure 1). SNVs may affect exon splicing or transcription [3], and are found more frequently than other types of genetic variations, such as differences in copy number, insertions, deletions, duplications, and rearrangements. SNVs in protein-coding regions have received the most attention, in spite of the fact that those regions account for only about 2% of the total human genome [7]. SNVs in the coding region can be synonymous if no amino acid change is produced, or non-synonymous if the substitution leads to a change in the protein sequence. The non-synonymous variants can be further divided into two categories: missense mutations, which lead to single amino acid changes, or nonsense mutations, which produce truncated or longer proteins (Figure 1).

International Journal of
Molecular Sciences

MDPI

Article

A Glimpse into the Structural Properties of the Intermediate and Transition State in the Folding of Bromodomain 2 Domain 2 by Φ Value Analysis

Leonore Novak [1,†], Maria Petrosino [1,†], Daniele Santorelli [1,†], Roberta Chiaraluce [1], Valerio Consalvi [1], Alessandra Pasquo [2] and Carlo Travaglini-Allocatelli [1,*]

[1] Dipartimento di Scienze Biochimiche "A. Rossi Fanelli", Sapienza University of Rome, 00185 Roma, Italy; leonore.novak@uniroma1.it (L.N.); maria.petrosino@uniroma1.it (M.P.); daniele.santorelli@uniroma1.it (D.S.); roberta.chiaraluce@uniroma1.it (R.C.); valerio.consalvi@uniroma1.it (V.C.)
[2] ENEA CR Frascati, Diagnostics and Metrology Laboratory FSN-TECFIS-DIM, 00044 Roma, Italy; alessandra.pasquo@enea.it
* Correspondence: carlo.travaglini@uniroma1.it
† These authors contributed equally to this work.

Abstract: Bromodomains (BRDs) are small protein interaction modules of about 110 amino acids that selectively recognize acetylated lysine in histones and other proteins. These domains have been identified in a variety of multi-domain proteins involved in transcriptional regulation or chromatin remodeling in eukaryotic cells. BRD inhibition is considered an attractive therapeutic approach in epigenetic disorders, particularly in oncology. Here, we present a Φ value analysis to investigate the folding pathway of the second domain of BRD2 (BRD2(2)). Using an extensive mutational analysis based on 25 site-directed mutants, we provide structural information on both the intermediate and late transition state of BRD2(2). The data reveal that the C-terminal region represents part of the initial folding nucleus, while the N-terminal region of the domain consolidates its structure only later in the folding process. Furthermore, only a small number of native-like interactions have been identified, suggesting the presence of a non-compact, partially folded state with scarce native-like characteristics. Taken together, these results indicate that, in BRD2(2), a hierarchical mechanism of protein folding can be described with non-native interactions that play a significant role in folding.

Keywords: bromodomain; protein folding; Φ value analysis; protein stability; mutagenesis; folding kinetics

Citation: Novak, L.; Petrosino, M.; Santorelli, D.; Chiaraluce, R.; Consalvi, V.; Pasquo, A.; Travaglini-Allocatelli, C. A Glimpse into the Structural Properties of the Intermediate and Transition State in the Folding of Bromodomain 2 Domain 2 by Φ Value Analysis. *Int. J. Mol. Sci.* **2021**, *22*, 5953. https://doi.org/10.3390/ijms22115953

Academic Editor: Yuri Lyubchenko

Received: 3 May 2021
Accepted: 27 May 2021
Published: 31 May 2021

1. Introduction

The correct folding process of the biological macromolecules is crucial for living cells, as their biochemical processes rely on finely tuned inter-molecular recognition events, which depend on structural complementarity between interacting molecules. This biochemical principle, known as the structure–function relationship, is particularly evident in the case of structurally complex macromolecules, such as proteins. However, notwithstanding decades of experimental, theoretical, and computational efforts, the mechanism of protein folding is still one of the major problems in molecular biology.

As for any chemical reaction, a clear description of the folding of a protein would require the identification and structural characterization of each of the molecular species transiently populated during the process [1]; however, in the case of protein folding, experimental difficulties arise because of the intrinsic cooperativity of the process and the large number of weak interactions forming from the denatured state to the native state in a biologically relevant timescale [2].

In this context, the description of the folding mechanism of proteins populating partially folded intermediates is particularly valuable, as it may offer the opportunity to follow the evolution of structure formation. In this context, the BET (bromo-extra-terminal domain) bromodomains (BRDs) represent a useful experimental system not only because

Common acronyms and abbreviations

AU	Absorbance Unit
ATP	Adenosine triphosphate
BSA	Bovine serum albumin
CD	Circular dichroism
CD-site	Common docking site
COSMIC	Catalogue of Somatic Mutation In Cancer
DNA	DeoxyriboNucleic Acid
DEAE	Diethylaminoethyl cellulose
dbSNP	Single Nucleotide Polymorphism Database
DSF	Differential Scanning Fluorimetry
ΔG	ΔGGibbs free energy
DTT	Dithiothreitol
$E_a^{a‡}$	Activation energy
EDTA	Ethylenediaminetetraacetic acid
ERK1/2	Extracellular signal–regulated kinases ½
GTP	Guanosine-5'-triphosphate
HapMap	The International HapMap Project
IPTG	Isopropil-β-D-1-tiogalattopiranoside
ITC	Isothermal titration calorimetry
LB	Lysogeny broth
MATLAB	Matrix Laboratory
MEK	Mitogen-activated protein kinase kinase
MES	2-(N-morpholino)ethanesulfonic acid
Mg(II)	Magnesium (II) ion

MgCl2	Magnesium Chloride
MgSo4	Magnesium Sulfate
MgATP	Adenosine 5′-triphosphate magnesium salt
NaCl	Sodium chloride
NCI	National Cancer Institute
NHGRI	National Human Genome Research Institute
nsSNP	non synonimus single nucleotide variants
NP-	unphosphorylated
NTS	Nuclear Traslocation signal
nsSNV	Non synonymous Single Nucleotide Variant
NTR	N-terminal region
OD600	Optical density
OMIM	Online Mendelian Inheritance in Man
P-	Phosphorylated
PDB	Protein Data Bank
RAS	Rat sarcoma virus
RAF	Proto-oncogene serine/threonine-protein kinase
SAV	Single Amino acid Variation
SNP	Single Nucleotide Polymorphism
SNV	Single Nucleotide Variant
SVD	Singular Value Decomposition
TARGET	Therapeutically Applicable Research to Generate Effective Treatments
TCGA	The Cancer Genome ATLAS
TCEP	Tris(2-carboxyethyl)phosphine hydrochloride
TEV	Tobacco etch virus

TGFβ	Transforming growth factor beta
T_m	Melting Temperature
Tris/HCl	Tris(hydroxymethyl)aminomethane hydrochloride
TXY	Threonine, amminoacid, Tyrosine
UniProt	Universal Protein

References

Adzhubei, Ivan A, Steffen Schmidt, Leonid Peshkin, Vasily E Ramensky, Anna Gerasimova, Peer Bork, Alexey S Kondrashov, e Shamil R Sunyaev. «A Method and Server for Predicting Damaging Missense Mutations». Nature Methods 7, n. 4 (2010): 248–49. https://doi.org/10.1038/nmeth0410-248.

Hassani-Allali, Abdellah, Gregory A. Wasney, Irene Chau, Bum Soo Hong, Guillermo Senisterra, Peter Loppnau, Zhen Shi, et al. «A Survey of Proteins Encoded by Non-Synonymous Single Nucleotide Polymorphisms Reveals a Significant Fraction with Altered Stability and Activity». Biochemical Journal 424, n. 1 (2009): 15–26. https://doi.org/10.1042/BJ20090723.

Ancien, F.; Pucci, F.; Godfroid, M.; Rooman, M. Prediction and Interpretation of Deleterious Coding Variants in Terms of Protein Structural Stability. Sci. Rep. n. 8 (2018) 4480. doi:10.1038/s41598-018-22531-2

Aronov A.M., Christopher Baker, Guy W. Bemis, Jingrong Cao, Guanjing Chen, Pamella J. Ford, Ursula A. Germann, et al. «Flipped Out: Structure-Guided Design of Selective Pyrazolylpyrrole ERK Inhibitors». Journal of Medicinal Chemistry 50, n. 6 (2007): 1280–87. https://doi.org/10.1021/jm061381f.

Arteaga Carlos L. and Jeffrey A. Engelman. «ERBB Receptors: From Oncogene Discovery to Basic Science to Mechanism-Based Cancer Therapeutics». Cancer Cell 25, n. 3 (2014): 282–303. https://doi.org/10.1016/j.ccr.2014.02.025

Bachegowda L., Kerry Morrone, Shannon L. Winski, Ioannis Mantzaris, Matthias Bartenstein, Nandini Ramachandra, Orsi Giricz, et al. “Pexmetinib:

A Novel Dual Inhibitor of Tie2 and p38 MAPK with Efficacy in Preclinical Models of Myelodysplastic Syndromes and Acute Myeloid Leukemia." Cancer Research 76, no. 16 (2016): 4841–49. https://doi.org/10.1158/0008-5472.CAN-15-3062.

Bailey M.H., Tokheim, C., Porta-Pardo E., Sengupta S., Bertrand D., Weerasinghe A., Colaprico A., Wendl M.C., Kim J., Reardon B., et al. Comprehensive Characterization of Cancer Driver Genes and Mutations. Cell 17 (2018): 371–385.e18, doi:10.1016/j.cell.2018.02.060.

Balu Kamaraj and Rituraj Purohit. "Mutational Analysis of TYR Gene and Its Structural Consequences in OCA1A." Gene 513, no. 1 (2013): 184–95. https://doi.org/10.1016/j.gene.2012.09.128.

Balu Kamaraj and Rituraj Purohit. «Computational Screening of Disease-Associated Mutations in OCA2 Gene». Cell Biochemistry and Biophysics 68, n. 1 (2014): 97–109. https://doi.org/10.1007/s12013-013-9697-2.

Balu Kamaraj and Rituraj Purohit. «In Silico Screening and Molecular Dynamics Simulation of Disease-Associated NsSNP in TYRP1 Gene and Its Structural Consequences in OCA3». BioMed Research International (2013): 1–13. https://doi.org/10.1155/2013/697051.

Bhagwat S.V., William T. McMillen, Shufen Cai, Baohui Zhao, Matthew Whitesell, Weihua Shen, Lisa Kindler, et al. "ERK Inhibitor LY3214996 Targets ERK Pathway–Driven Cancers: A Therapeutic Approach Toward Precision Medicine." Molecular Cancer Therapeutics 19, no. 2 (2020): 325–36. https://doi.org/10.1158/1535-7163.MCT-19-0183.

Boga S. B., Yongqi Deng, Liang Zhu, Yang Nan, Alan B. Cooper, Gerald W. Shipps, Ronald Doll, et al. "MK-8353: Discovery of an Orally Bioavailable

Dual Mechanism ERK Inhibitor for Oncology." ACS Medicinal Chemistry Letters 9, no. 7 (2018): 761–67. https://doi.org/10.1021/acsmedchemlett.8b00220.

Balu K., e Rituraj Purohit. «In Silico Screening and Molecular Dynamics Simulation of Disease-Associated NsSNP in TYRP1 Gene and Its Structural Consequences in OCA3». BioMed Research International 2013 (2013): 1–13. https://doi.org/10.1155/2013/697051.

Bandaru P., Kondo Y. and Kuriyan J. «The Interdependent Activation of Son-of-Sevenless and Ras». Cold Spring Harbor Perspectives in Medicine 9, n. 2 (2019): a031534. https://doi.org/10.1101/cshperspect.a031534.

Benjwal S. «Monitoring Protein Aggregation during Thermal Unfolding in Circular Dichroism Experiments». Protein Science 15, n. 3 (2006): 635–39. https://doi.org/10.1110/ps.051917406.

Berti Denise A. and Rony Seger. "The Nuclear Translocation of ERK." In ERK Signaling, edited by Gerardo Jimenez, 1487:175–94. New York, NY: Springer New York (2017). http://link.springer.com/10.1007/978-1-4939-6424-6_13.

Buniello A., MacArthur, J.A.L., Cerezo, M., Harris, L.W., Hayhurst, J., Malangone, C., McMahon, A., Morales, J., Mountjoy, E., Sollis, E., et al. The NHGRI-EBI GWAS Catalog of Published Genome-Wide Association Studies, Targeted Arrays and Summary Statistics (2019). Nucleic Acids Res. 2019, 47, D1005–D1012, doi:10.1093/nar/gky1120.

Burrows F., Linda Kessler, Jeffrey Chen, Xin Gao, Rasmus Hansen, Shuangwei Li, Carol Thach, et al. "Abstract 5168: KO-947, a Potent ERK Inhibitor with Robust Preclinical Single Agent Activity in MAPK Pathway

Dysregulated Tumors," 5168–5168. American Association for Cancer Research, (2017). https://doi.org/10.1158/1538-7445.AM2017-5168.

Cairns John. "Mutation Selection and the Natural History of Cancer." Nature 255, no. 5505 (1975): 197–200. https://doi.org/10.1038/255197a0.

Callaway K., William F. Waas, Mark A. Rainey, Pengyu Ren, e Kevin N. Dalby. «Phosphorylation of the Transcription Factor Ets-1 by ERK2: Rapid Dissociation of ADP and Phospho-Ets-1». Biochemistry 49, n. 17 (2010): 3619–30. https://doi.org/10.1021/bi100199q.

Capriotti, E., P. Fariselli, R. Calabrese, e R. Casadio. «Predicting Protein Stability Changes from Sequences Using Support Vector Machines». Bioinformatics 21, n. Suppl 2 (2005): ii54–58. https://doi.org/10.1093/bioinformatics/bti1109.

Carles F., Stéphane Bourg, Christophe Meyer, and Pascal Bonnet. "PKIDB: A Curated, Annotated and Updated Database of Protein Kinase Inhibitors in Clinical Trials." Molecules 23, no. 4 (2018): 908. https://doi.org/10.3390/molecules23040908.

Casadio R., Marco Vassura, Shalinee Tiwari, Piero Fariselli, e Pier Luigi Martelli. «Correlating Disease-Related Mutations to Their Effect on Protein Stability: A Large-Scale Analysis of the Human Proteome». Human Mutation 32, n. 10 (2011): 1161–70. https://doi.org/10.1002/humu.21555.

Chaikuad Apirat, Eliana M C Tacconi, Jutta Zimmer, Yanke Liang, Nathanael S Gray, Madalena Tarsounas, e Stefan Knapp. «A Unique Inhibitor Binding Site in ERK1/2 Is Associated with Slow Binding Kinetics». Nature Chemical Biology 10, n. 10 (2014): 853–60. https://doi.org/10.1038/nchembio.1629.

Chan D.W., Vincent W.S. Liu, George S.W. Tsao, Kwok-Ming Yao, Toru Furukawa, Karen K.L. Chan, and Hextan Y.S. Ngan. “Loss of MKP3 Mediated by Oxidative Stress Enhances Tumorigenicity and Chemoresistance of Ovarian Cancer Cells.” Carcinogenesis 29, no. 9 (2008): 1742–50. https://doi.org/10.1093/carcin/bgn167.

Chan S.H., Weng Khong Lim, Nur Diana Binte Ishak, Shao-Tzu Li, Wei Lin Goh, Gek San Tan, Kiat Hon Lim, et al. “Germline Mutations in Cancer Predisposition Genes Are Frequent in Sporadic Sarcomas.” Scientific Reports 7, no. 1 (2017). https://doi.org/10.1038/s41598-017-10333-x.

Cheng Heung-Chin, Robert Z. Qi, Hemant Paudel, and Hong-Jian Zhu. “Regulation and Function of Protein Kinases and Phosphatases.” Enzyme Research 2011 (2011): 1–3. https://doi.org/10.4061/2011/794089.

Cheng P. F., Reinhard Dummer, e Mitchell P Levesque. «Data mining the cancer genome atlas in the era of precision cancer medicine», (2015). https://doi.org/10.5167/UZH-113092.

Cheng T.M. K., Yu-En Lu, Michele Vendruscolo, Pietro Lio’and Tom L. Blundell. «Prediction by Graph Theoretic Measures of Structural Effects in Proteins Arising from Non-Synonymous Single Nucleotide Polymorphisms». A cura di Ruth Nussinov. PLoS Computational Biology 4, n. 7 (2008): e1000135. https://doi.org/10.1371/journal.pcbi.1000135.

Collins FS, Guyer MS & Chakravarti A. Variations on a theme: cataloging human DNA sequence variation. Science 278, (1997): 1580−1581.

Cristofaro R., Andrea Carotti, Sepideh Akhavan, Roberta Palla, Flora Peyvandi, Cosimo Altomare, e Pier Mannuccio Mannucci. «The Natural Mutation by Deletion of Lys9 in the Thrombin A-Chain Affects the PKa Value

of Catalytic Residues, the Overall Enzyme's Stability and Conformational Transitions Linked to Na+ Binding». FEBS Journal 273, n. 1 (2006): 159–69. https://doi.org/10.1111/j.1742-4658.2005.05052.x.

Dar Arvin C. and Kevan M. Shokat. «The Evolution of Protein Kinase Inhibitors from Antagonists to Agonists of Cellular Signaling». Annual Review of Biochemistry 80, n. 1 (2011): 769–95. https://doi.org/10.1146/annurev-biochem-090308-173656.

Dehouck Y., Jean Marc Kwasigroch, Dimitri Gilis and Marianne Rooman. «PoPMuSiC 2.1: A Web Server for the Estimation of Protein Stability Changes upon Mutation and Sequence Optimality». BMC Bioinformatics 12, n. 1 (2011): 151. https://doi.org/10.1186/1471-2105-12-151.

Dohlman H. G. and Sharon L. Campbell. «Regulation of Large and Small G Proteins by Ubiquitination». Journal of Biological Chemistry 294, n. 49 (2019): 18613–23. https://doi.org/10.1074/jbc.REV119.011068.

Eglen R. and Terry Reisine. «Drug Discovery and the Human Kinome: Recent Trends». Pharmacology & Therapeutics 130, n. 2 (2011): 144–56. https://doi.org/10.1016/j.pharmthera.2011.01.007.

Fantz D.A., Dave Jacobs, Danielle Glossip, e Kerry Kornfeld. «Docking Sites on Substrate Proteins Direct Extracellular Signal-Regulated Kinase to Phosphorylate Specific Residues». Journal of Biological Chemistry 276, n. 29 (2001): 27256–65. https://doi.org/10.1074/jbc.M102512200.

Fariselli P., Pier Luigi Martelli, Castrense Savojardo, e Rita Casadio. «INPS: Predicting the Impact of Non-Synonymous Variations on Protein Stability from Sequence». Bioinformatics 31, n. 17 (2015): 2816–21. https://doi.org/10.1093/bioinformatics/btv291.

Flannick J. and Florez J.C. Type 2 Diabetes: Genetic Data Sharing to Advance Complex Disease Research. Nat. Rev. Genet. n. 17, (2016): 535–549, doi:10.1038/nrg.2016.56.

Forbes S. A., N. Bindal, S. Bamford, C. Cole, C. Y. Kok, D. Beare, M. Jia, et al. «COSMIC: Mining Complete Cancer Genomes in the Catalogue of Somatic Mutations in Cancer». Nucleic Acids Research 39, n. Database (2011): D945–50. https://doi.org/10.1093/nar/gkq929.

Furukawa T., Makoto Sunamura, Fuyuhiko Motoi, Seiki Matsuno, and Akira Horii. "Potential Tumor Suppressive Pathway Involving DUSP6/MKP-3 in Pancreatic Cancer." The American Journal of Pathology 162, no. 6 (2003): 1807–15. https://doi.org/10.1016/S0002-9440(10)64315-5.

Gan K. A., Sebastian Carrasco Pro, Jared A. Sewell, and Juan I. Fuxman Bass. "Identification of Single Nucleotide Non-Coding Driver Mutations in Cancer." Frontiers in Genetics 9 (2018). https://doi.org/10.3389/fgene.2018.00016.

Gao K., Rick Oerlemans and Matthew R. Groves. "Theory and Applications of Differential Scanning Fluorimetry in Early-Stage Drug Discovery." Biophysical Reviews 12, no. 1 (2020): 85–104. https://doi.org/10.1007/s12551-020-00619-2.

Garland W., Robert Benezra, e Jaideep Chaudhary. «Targeting Protein–Protein Interactions to Treat Cancer—Recent Progress and Future Directions». In Annual Reports in Medicinal Chemistry. n. 48(2013):227–45. Elsevier. https://doi.org/10.1016/B978-0-12-417150-3.00015-6.

González-Pérez Abel and Nuria López-Bigas. «Improving the Assessment of the Outcome of Nonsynonymous SNVs with a Consensus Deleteriousness

Score, Condel». The American Journal of Human Genetics 88, n. 4 (2011): 440–49. https://doi.org/10.1016/j.ajhg.2011.03.004.

Greaves Mel and Carlo C. Maley. “Clonal Evolution in Cancer.” Nature 481, no. 7381 (2012): 306–13. https://doi.org/10.1038/nature10762.

Guo Yan-Jun, Wei-Wei Pan, Sheng-Bing Liu, Zhong-Fei Shen, Ying Xu, e Ling-Ling Hu. «ERK/MAPK signalling pathway and tumorigenesis (Review)». Experimental and Therapeutic Medicine. (2020). https://doi.org/10.3892/etm.2020.8454.

Hamosh, A. «Online Mendelian Inheritance in Man (OMIM), a Knowledgebase of Human Genes and Genetic Disorders». Nucleic Acids Research 33, n. Database issue (2004): D514–17. https://doi.org/10.1093/nar/gki033.

Hancock C.N., Alba Macias, Eun Kyoung Lee, Su Yeon Yu, Alexander D. MacKerell and Paul Shapiro. «Identification of Novel Extracellular Signal-Regulated Kinase Docking Domain Inhibitors». Journal of Medicinal Chemistry 48, n. 14 (luglio 2005): 4586–95. https://doi.org/10.1021/jm0501174.

Hardcastle I.R. «Protein–Protein Interaction Inhibitors in Cancer». In Comprehensive Medicinal Chemistry III, 154–201. Elsevier, (2017). 10.1016/B978-0-12-409547-2.12392-3

Hecht M., Yana Bromberg and Burkhard Rost. «News from the Protein Mutability Landscape». Journal of Molecular Biology 425, n. 21 (2013): 3937–48. https://doi.org/10.1016/j.jmb.2013.07.028.

Holt L. J., B. B. Tuch, J. Villen, A. D. Johnson, S. P. Gygi, e D. O. Morgan. «Global Analysis of Cdk1 Substrate Phosphorylation Sites Provides Insights into Evolution». Science 325, n. 5948 (2009): 1682–86. https://doi.org/10.1126/science.1172867.

Hu C., Steven N. Hart, Eric C. Polley, Rohan Gnanaolivu, Hermela Shimelis, Kun Y. Lee, Jenna Lilyquist, et al. "Association Between Inherited Germline Mutations in Cancer Predisposition Genes and Risk of Pancreatic Cancer." JAMA 319, no. 23 (2018): 2401. https://doi.org/10.1001/jama.2018.6228.

Hua Rui-Xi, He-Ping Li, Yan-Bing Liang, Jin-Hong Zhu, Bing Zhang, Sheng Ye, Qiang-Sheng Dai, Shi-Qiu Xiong, Yong Gu, e Xiang-Zhou Sun. «Association between the PARP1 Val762Ala Polymorphism and Cancer Risk: Evidence from 43 Studies». A cura di Balraj Mittal. PLoS ONE 9, n. 1 (2014): e87057. https://doi.org/10.1371/journal.pone.0087057.

Hunter Tony. "Protein Kinases and Phosphatases: The Yin and Yang of Protein Phosphorylation and Signaling." Cell 80, no. 2 (1995): 225–36. https://doi.org/10.1016/0092-8674(95)90405-0.

Hunter Tony. «A Thousand and One Protein Kinases». Cell 50, n. 6 (1987): 823–29. https://doi.org/10.1016/0092-8674(87)90509-5.

Hurst J.M., Lisa E.M. McMillan, Craig T. Porter, James Allen, Adebola Fakorede, e Andrew C.R. Martin. «The SAAPdb Web Resource: A Large-Scale Structural Analysis of Mutant Proteins». Human Mutation 30, n. 4 (2009): 616–24. https://doi.org/10.1002/humu.20898.

Isom D. G., Brian R. Cannon, Carlos A. Castañeda, Aaron Robinson, and Bertrand García-Moreno E. "High Tolerance for Ionizable Residues in the Hydrophobic Interior of Proteins." Proceedings of the National Academy of

Sciences 105, no. 46 (2008): 17784–88. https://doi.org/10.1073/pnas.0805113105.

Jasek-Gajda E., Halina Jurkowska, Małgorzata Jasińska, Jan A. Litwin, and Grzegorz J. Lis. "Combination of ERK2 Inhibitor VX-11e and Voreloxin Synergistically Enhances Anti-Proliferative and pro-Apoptotic Effects in Leukemia Cells." Apoptosis 24, no. 11–12 (2019): 849–61. https://doi.org/10.1007/s10495-019-01564-6.

Jacobs D., Danielle Glossip, Heming Xing, Anthony J. Muslin, e Kerry Kornfeld. «Multiple docking sites on substrate proteins form a modular system that mediates recognition by ERK MAP kinase», (1999), Genes Dev. edizione, par. 13(2).

Jemal A., Freddie Bray, Melissa M. Center, Jacques Ferlay, Elizabeth Ward and David Forman. «Global Cancer Statistics». CA: A Cancer Journal for Clinicians 61, n. 2 (2011): 69–90. https://doi.org/10.3322/caac.20107.

Jones J. C., Lindsay A. Renfro, Humaid O. Al-Shamsi, Alexa B. Schrock, Andrew Rankin, Ben Y. Zhang, Pashtoon M. Kasi, et al. «Non-V600 BRAF Mutations Define a Clinically Distinct Molecular Subtype of Metastatic Colorectal Cancer». Journal of Clinical Oncology 35, n. 23 (2017): 2624–30. https://doi.org/10.1200/JCO.2016.71.4394.

Juritz E., Maria Fornasari, Pier Martelli, Piero Fariselli, Rita Casadio and Gustavo Parisi. «On the Effect of Protein Conformation Diversity in Discriminating among Neutral and Disease Related Single Amino Acid Substitutions». BMC Genomics 13, n. Suppl 4 (2012): S5. https://doi.org/10.1186/1471-2164-13-S4-S5.

Karchin R., M. Diekhans, L. Kelly, D. J. Thomas, U. Pieper, N. Eswar, D. Haussler, e A. Sali. «LS-SNP: Large-Scale Annotation of Coding Non-Synonymous SNPs Based on Multiple Information Sources». Bioinformatics 21, n. 12 (2005): 2814–20. https://doi.org/10.1093/bioinformatics/bti442.

Kidger A.M., James Sipthorp and Simon J. Cook. «ERK1/2 Inhibitors: New Weapons to Inhibit the RAS-Regulated RAF-MEK1/2-ERK1/2 Pathway». Pharmacology & Therapeutics 187 (2018): 45–60. https://doi.org/10.1016/j.pharmthera.2018.02.007.

Knighton D., J. Zheng, L. Ten Eyck, V. Ashford, N. Xuong, S. Taylor, e J. Sowadski. «Crystal Structure of the Catalytic Subunit of Cyclic Adenosine Monophosphate-Dependent Protein Kinase». Science 253, n. 5018 (1991): 407–14. https://doi.org/10.1126/science.1862342.

Kornev Alexandr P.and Susan S. Taylor. «Defining the Conserved Internal Architecture of a Protein Kinase». Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics 1804, n. 3 (2010): 440–44. https://doi.org/10.1016/j.bbapap.2009.10.017.

Koukouritaki S.B., Mark T. Poch, Marilyn C. Henderson, Lisbeth K. Siddens, Sharon K. Krueger, Jonathan E. VanDyke, David E. Williams, Nicholas M. Pajewski, Tao Wang, e Ronald N. Hines. «Identification and Functional Analysis of Common Human Flavin-Containing Monooxygenase 3 Genetic Variants». Journal of Pharmacology and Experimental Therapeutics 320, n. 1 (2007): 266–73. https://doi.org/10.1124/jpet.106.112268.

Kucukkal T.G., Marharyta Petukh, Lin Li, e Emil Alexov. «Structural and Physico-Chemical Effects of Disease and Non-Disease NsSNPs on Proteins».

Current Opinion in Structural Biology 32 (2015): 18–24. https://doi.org/10.1016/j.sbi.2015.01.003.

Kucukkal T., Ye Yang, Susan Chapman, Weiguo Cao, e Emil Alexov. «Computational and Experimental Approaches to Reveal the Effects of Single Nucleotide Polymorphisms with Respect to Disease Diagnostics». International Journal of Molecular Sciences 15, n. 6 (2014): 9670–9717. https://doi.org/10.3390/ijms15069670.

Kumar A., Vidya Rajendran, Rao Sethumadhavan, Rituraj Purohit. «Evidence of Colorectal Cancer-Associated Mutation in MCAK: A Computational Report». Cell Biochemistry and Biophysics 67, n. 3 (2013): 837–51. https://doi.org/10.1007/s12013-013-9572-1.

Kumar A., Vidya Rajendran, Rao Sethumadhavan, Rituraj Purohit. «Roadmap to Determine the Point Mutations Involved in Cardiomyopathy Disorder: A Bayesian Approach». Gene 519, n. 1 (2013): 34–40. https://doi.org/10.1016/j.gene.2013.01.056.

Lappalainen T.; Scott, A.J.; Brandt, M.; Hall, I.M. Genomic Analysis in the Age of Human Genome Sequencing. Cell 177 (2019): 70–84, doi:10.1016/j.cell.2019.02.032.

Larsen C.P., Thomas C. Pearson, Andrew B. Adams, Paul Tso, Nozomu Shirasugi, Elizabeth StrobertM, Dan Anderson, et al. «Rational Development of LEA29Y (Belatacept), a High-Affinity Variant of CTLA4-Ig with Potent Immunosuppressive Properties». American Journal of Transplantation 5, n. 3 (2005): 443–53. https://doi.org/10.1111/j.1600-6143.2005.00749.x.

Leaver-Fay A., Michael Tyka, Steven M. Lewis, Oliver F. Lange, James Thompson, Ron Jacak, Kristian W. Kaufman, et al. «Rosetta3». In Methods

in Enzymology, 487(2011):545–74. Elsevier.https://doi.org/10.1016/B978-0-12-381270-4.00019-6.

Lee B., Ben Tran, Arthur L. Hsu, Graham R. Taylor, Stephen B. Fox, Andrew Fellowes, Renata Marquis, et al. "Exploring the Feasibility and Utility of Exome-Scale Tumour Sequencing in a Clinical Setting: Feasibility of Clinical Exome Sequencing." Internal Medicine Journal 48, no. 7 (2018): 786–94. https://doi.org/10.1111/imj.13806.

Lee T., Andrew N Hoofnagle, Yukihito Kabuyama, James Stroud, Xiaoshan Min, Elizabeth J Goldsmith, Lin Chen, Katheryn A Resing, e Natalie G Ahn. «Docking Motif Interactions in MAP Kinases Revealed by Hydrogen Exchange Mass Spectrometry». Molecular Cell 14, n. 1 (2004): 43–55. https://doi.org/10.1016/S1097-2765(04)00161-3.

Levy S., Granger S., Pauline C Ng, Lars F., Aaron L H., Brian P W., Nelson A., et al. «The Diploid Genome Sequence of an Individual Human». A cura di Edward M Rubin. PLoS Biology 5, n. 10 (2007): e254. https://doi.org/10.1371/journal.pbio.0050254.

Lichtenstein P., Niels V. Holm, Pia K. Verkasalo, Anastasia Iliadou, Jaakko Kaprio, Markku Koskenvuo, Eero Pukkala, Axel Skytthe and Kari Hemminki. «Environmental and Heritable Factors in the Causation of Cancer — Analyses of Cohorts of Twins from Sweden, Denmark, and Finland». New England Journal of Medicine 343, n. 2 (2000): 78–85. https://doi.org/10.1056/NEJM200007133430201.

Lim J., Elizabeth H. Kelley, Joey L. Methot, Hua Zhou, Alessia Petrocchi, Hongmin Chen, Susan E. Hill, et al. «Discovery of 1-(1 H -Pyrazolo[4,3- c]Pyridin-6-Yl)Urea Inhibitors of Extracellular Signal-Regulated Kinase

(ERK) for the Treatment of Cancers». Journal of Medicinal Chemistry 59, n. 13 (2016): 6501–11. https://doi.org/10.1021/acs.jmedchem.6b00708.

Lin William and Un Jung Kang. «Characterization of PINK1 Processing, Stability, and Subcellular Localization». Journal of Neurochemistry 106, n. 1 (2008): 464–74. https://doi.org/10.1111/j.1471-4159.2008.05398.x.

Liu Y., Easton J., Shao Y.; Maciaszek J., Wang Z., Wilkinson M.R., McCastlain K., Edmonson M., Pounds S.B., Shi L., et al. The Genomic Landscape of Pediatric and Young Adult T-Lineage Acute Lymphoblastic Leukemia. Nat. Genet. 49 (2017):1211–1218, doi:10.1038/ng.3909.

Liu Q., Yogesh Sabnis, Zheng Zhao, Tinghu Zhang, Sara J. Buhrlage, Lyn H. Jones and Nathanael S. Gray. «Developing Irreversible Inhibitors of the Protein Kinase Cysteinome». Chemistry & Biology 20, n. 2 (2013): 146–59. https://doi.org/10.1016/j.chembiol.2012.12.006.

Liu, Xiaoming, Xueqiu Jian and Eric Boerwinkle. «DbNSFP v2.0: A Database of Human Non-Synonymous SNVs and Their Functional Predictions and Annotations». Human Mutation 34, n. 9 (2013): E2393–2402. https://doi.org/10.1002/humu.22376.

Lloyd Alison C. «Distinct functions for ERKs?» Journal of Biology 5, n. 5 (2006): 13. https://doi.org/10.1186/jbiol46.

Lori C., Antonella Lantella, Alessandra Pasquo, Leila T. Alexander, Stefan Knapp, Roberta Chiaraluce and Valerio Consalvi. «Effect of Single Amino Acid Substitution Observed in Cancer on Pim-1 Kinase Thermodynamic Stability and Structure». A cura di Annalisa Pastore. PLoS ONE 8, n. 6 (2013): e64824. https://doi.org/10.1371/journal.pone.0064824.

MacArthur D. G., S. Balasubramanian, A. Frankish, N. Huang, J. Morris, K. Walter, L. Jostins, et al. «A Systematic Survey of Loss-of-Function Variants in Human Protein-Coding Genes». Science 335, n. 6070 (2012): 823–28. https://doi.org/10.1126/science.1215040.

Malhotra S.; Alsulami, A.F.; Heiyun, Y.; Ochoa, B.M.; Jubb, H.; Forbes, S.; Blundell, T.L. Understanding the Impacts of Missense Mutations on Structures and Functions of Human Cancer-Related Genes: A Preliminary Computational Analysis of the COSMIC Cancer Gene Census. PLoS ONE 2019, 14, e0219935, doi:10.1371/journal.pone.0219935.

Manning G. «The Protein Kinase Complement of the Human Genome». Science 298, n. 5600 (2002): 1912–34. https://doi.org/10.1126/science.1075762.

Milanese Jean-Sébastien and Edwin Wang. "Germline Mutations and Their Clinical Applications in Cancer." Breast Cancer Management 8, no. 1 (2019): BMT23. https://doi.org/10.2217/bmt-2019-0005.

Morais V.A., Patrik Verstreken, Anne Roethig, Joél Smet, An Snellinx, Mieke Vanbrabant, Dominik Haddad, et al. «Parkinson's Disease Mutations in PINK1 Result in Decreased Complex I Activity and Deficient Synaptic Function». EMBO Molecular Medicine 1, n. 2 (2009): 99–111. https://doi.org/10.1002/emmm.200900006.

Morris E.J., Sharda Jha, Clifford R. Restaino, Priya Dayananth, Hugh Zhu, Alan Cooper, Donna Carr, et al. «Discovery of a Novel ERK Inhibitor with Activity in Models of Acquired Resistance to BRAF and MEK Inhibitors».

Cancer Discovery 3, n. 7 (2013): 742–50. https://doi.org/10.1158/2159-8290.CD-13-0070.

Morrison D.K., Monica S. Murakami, and Vaughn Cleghon. "Protein Kinases and Phosphatases in the Drosophila Genome." Journal of Cell Biology 150, no. 2 (2000): F57–62. https://doi.org/10.1083/jcb.150.2.F57.

Muñoz-Maldonado C., Zimmer Y and Medová M. «A Comparative Analysis of Individual RAS Mutations in Cancer Biology». Frontiers in Oncology 9 (2019): 1088. https://doi.org/10.3389/fonc.2019.01088.

Muta Y., Matsuda M. and Imajo M. «Divergent Dynamics and Functions of ERK MAP Kinase Signaling in Development, Homeostasis and Cancer: Lessons from Fluorescent Bioimaging». Cancers 11, n. 4 (2019): 513. https://doi.org/10.3390/cancers11040513.

Nakao Hiroyuki, Chihiro Hayashi, Keisuke Ikeda, Hiroaki Saito, Hidemi Nagao, and Minoru Nakano. "Effects of Hydrophilic Residues and Hydrophobic Length on Flip-Flop Promotion by Transmembrane Peptides." The Journal of Physical Chemistry B 122, no. 15 (2018): 4318–24. https://doi.org/10.1021/acs.jpcb.8b00298.

Ng P.C. and Steven Henikoff. «Predicting the Effects of Amino Acid Substitutions on Protein Function». Annual Review of Genomics and Human Genetics 7, n. 1 (2006): 61–80. https://doi.org/10.1146/annurev.genom.7.080505.115630.

Nowell P.C. The Clonal Evolution of Tumor Cell Populations. Science. 194, (1976): 23–28.

Nuytemans K., Jessie Theuns, Marc Cruts and Christine Van Broeckhoven. «Genetic Etiology of Parkinson Disease Associated with Mutations in the SNCA, PARK2, PINK1, PARK7, and LRRK2 Genes: A Mutation Update». Human Mutation 31, n. 7 (2010): 763–80. https://doi.org/10.1002/humu.21277.

Ode Hirotaka, Shou Matsuyama, Masayuki Hata, Saburo Neya, Junko Kakizawa, Wataru Sugiura and Tyuji Hoshino. «Computational Characterization of Structural Role of the Non-Active Site Mutation M36I of Human Immunodeficiency Virus Type 1 Protease». Journal of Molecular Biology 370, n. 3 (2007): 598–607. https://doi.org/10.1016/j.jmb.2007.04.081.

Parthiban V., M. M. Gromiha, e D. Schomburg. «CUPSAT: Prediction of Protein Stability upon Point Mutations». Nucleic Acids Research 34, n. Web Server (2006): W239–42. https://doi.org/10.1093/nar/gkl190.

Petrosino M., Laura Lori, Alessandra Pasquo, Clorinda Lori, Valerio Consalvi, Velia Minicozzi, Silvia Morante, et al. «Single-Nucleotide Polymorphism of PPARγ, a Protein at the Crossroads of Physiological and Pathological Processes». International Journal of Molecular Sciences 18, n. 2 (2017): 361. https://doi.org/10.3390/ijms18020361.

Petukh M., Tugba G. Kucukkal, e Emil Alexov. «On Human Disease-Causing Amino Acid Variants: Statistical Study of Sequence and Structural Patterns». Human Mutation 36, n. 5 (2015): 524–34. https://doi.org/10.1002/humu.22770.

Rainey M.A., Kari Callaway, Richard Barnes, Brian Wilson, e Kevin N. Dalby. «Proximity-Induced Catalysis by the Protein Kinase ERK2». Journal

of the American Chemical Society 127, n. 30 (2005): 10494–95. https://doi.org/10.1021/ja052915p

Ramensky V. «Human non-synonymous SNPs: server and survey». Nucleic Acids Research 30, n. 17 (2002): 3894–3900. https://doi.org/10.1093/nar/gkf493.

Roskoski Robert. «A Historical Overview of Protein Kinases and Their Targeted Small Molecule Inhibitors». Pharmacological Research 100 (2015): 1–23. https://doi.org/10.1016/j.phrs.2015.07.010.

Roskoski Robert. «ERK1/2 MAP Kinases: Structure, Function, and Regulation». Pharmacological Research 66, n. 2 (2012): 105–43. https://doi.org/10.1016/j.phrs.2012.04.005.

Royer C.A., Craig J. Mann, e C. Robert Matthews. «Resolution of the Fluorescence Equilibrium Unfolding Profile of Trp Aporepressor Using Single Tryptophan Mutants». Protein Science 2, n. 11 (1993): 1844–52. https://doi.org/10.1002/pro.5560021106.

Sammons R.M., Nicole A. Perry, Yangmei Li, Eun Jeong Cho, Andrea Piserchio, Diana P. Zamora-Olivares, Ranajeet Ghose, et al. «A Novel Class of Common Docking Domain Inhibitors That Prevent ERK2 Activation and Substrate Phosphorylation». ACS Chemical Biology 14, n. 6 (2019): 1183–94. https://doi.org/10.1021/acschembio.9b00093.

Santoro M.M. and D. W. Bolen. «Unfolding Free Energy Changes Determined by the Linear Extrapolation Method. 1. Unfolding of Phenylmethanesulfonyl .Alpha.-Chymotrypsin Using Different Denaturants». Biochemistry 27, n. 21 (1988): 8063–68. https://doi.org/10.1021/bi00421a014.

Schaefer, C., A. Meier, B. Rost, e Y. Bromberg. «SNPdbe: Constructing an NsSNP Functional Impacts Database». Bioinformatics 28, n. 4 (2012): 601–2. https://doi.org/10.1093/bioinformatics/btr705.

Schubbert S., Kevin Shannon, e Gideon Bollag. «Hyperactive Ras in Developmental Disorders and Cancer». Nature Reviews Cancer 7, n. 4 (2007): 295–308. https://doi.org/10.1038/nrc2109.

Schymkowitz J., J. Borg, F. Stricher, R. Nys, F. Rousseau, e L. Serrano. «The FoldX Web Server: An Online Force Field». Nucleic Acids Research 33, n. Web Server (2005): W382–88. https://doi.org/10.1093/nar/gki387.

Serber Zach and James E. Ferrell. «Tuning Bulk Electrostatics to Regulate Protein Function». Cell 128, n. 3 (2007): 441–44. https://doi.org/10.1016/j.cell.2007.01.018.

Sharrocks A.D, Shen-His Yang and Alex Galanis. «Docking Domains and Substrate-Specificity Determination for MAP Kinases». Trends in Biochemical Sciences 25, n. 9 (2000): 448–53. https://doi.org/10.1016/S0968-0004(00)01627-3.

Sheridan D.L., Yong Kong, Sirlester A. Parker, Kevin N. Dalby, e Benjamin E. Turk. «Substrate Discrimination among Mitogen-Activated Protein Kinases through Distinct Docking Sequence Motifs». Journal of Biological Chemistry 283, n. 28 (2008): 19511–20. https://doi.org/10.1074/jbc.M801074200.

Sherry S. T. «dbSNP: the NCBI database of genetic variation». Nucleic Acids Research 29, n. 1 (2001): 308–11. https://doi.org/10.1093/nar/29.1.308.

Shi Zhen and John Moult. «Structural and Functional Impact of Cancer-Related Missense Somatic Mutations». Journal of Molecular Biology 413, n. 2 (2011): 495–512. https://doi.org/10.1016/j.jmb.2011.06.046.

Singh V., Mahendra Ram, Rajesh Kumar, Raju Prasad, Birendra Kumar Roy, and Kaushal Kumar Singh. "Phosphorylation: Implications in Cancer." The Protein Journal 36, no. 1 (2017): 1–6. https://doi.org/10.1007/s10930-017-9696-z.

Shirley Bret A., Patrick Stanssens, Ulrich Hahn and C. Nick Pace. «Contribution of Hydrogen Bonding to the Conformational Stability of Ribonuclease T1». Biochemistry 31, n. 3 (1992): 725–32. https://doi.org/10.1021/bi00118a013.

Shults M.D., Dora Carrico-Moniz, e Barbara Imperiali. «Optimal Sox-Based Fluorescent Chemosensor Design for Serine/Threonine Protein Kinases». Analytical Biochemistry 352, n. 2 (2006): 198–207. https://doi.org/10.1016/j.ab.2006.03.003.

Sim, Ngak-Leng, Prateek Kumar, Jing Hu, Steven Henikoff, Georg Schneider, e Pauline C. Ng. «SIFT Web Server: Predicting Effects of Amino Acid Substitutions on Proteins». Nucleic Acids Research 40, n. W1 (2012): W452–57. https://doi.org/10.1093/nar/gks539.

Simanshu K.D., Nissley D. V. and McCormick F. «RAS Proteins and Their Regulators in Human Disease». Cell 170, n. 1 (2017): 17–33. https://doi.org/10.1016/j.cell.2017.06.009.

Simpson I., Mark J. Anderton, David M. Andrews, Jason Breed, Emma Davies, Judit E. Debreczeni, Vikki Flemington, et al. "Abstract 1647:

Discovery of AZD0364, a Potent and Selective Oral Inhibitor of ERK1/2 That Is Efficacious in Both Monotherapy and Combination Therapy in Models of NSCLC," (2018):1647–1647. American Association for Cancer Research. https://doi.org/10.1158/1538-7445.AM2018-1647.

Singh J., Russell C. Petter, Thomas A. Baillie and Adrian Whitty. «The Resurgence of Covalent Drugs». Nature Reviews Drug Discovery 10, n. 4 (aprile 2011): 307–17. https://doi.org/10.1038/nrd3410.

Sipthorp J., Honorine Lebraud, Rebecca Gilley, Andrew M. Kidger, Hanneke Okkenhaug, Marc Saba-El-Leil, Sylvain Meloche, Christopher J. Caunt, Simon J. Cook and Tom D. Heightman. «Visualization of Endogenous ERK1/2 in Cells with a Bioorthogonal Covalent Probe». Bioconjugate Chemistry 28, n. 6 (2017): 1677–83. https://doi.org/10.1021/acs.bioconjchem.7b00152.

Sreekanth G.P., Aporn Chuncharunee, Aunchalee Sirimontaporn, Jutatip Panaampon, Chatchawan Srisawat, Atthapan Morchang, Shilu Malakar, et al. "Role of ERK1/2 Signaling in Dengue Virus-Induced Liver Injury." Virus Research 188 (2014): 15–26. https://doi.org/10.1016/j.virusres.2014.03.025.

Stanley C.M., Shamil R. Sunyaev, Marc S. Greenblatt, e William S. Oetting. «Clinically Relevant Variants - Identifying, Collecting, Interpreting, and Disseminating: The 2013 Annual Scientific Meeting of the Human Genome Variation Society». Human Mutation 35, n. 4 (2014): 505–10. https://doi.org/10.1002/humu.22516.

Stefl Shannon, Hafumi Nishi, Marharyta Petukh, Anna R. Panchenko and Emil Alexov. «Molecular Mechanisms of Disease-Causing Missense

Mutations». Journal of Molecular Biology 425, n. 21 (2013): 3919–36. https://doi.org/10.1016/j.jmb.2013.07.014.

Stehr H., Seon-Hi J Jang, José M Duarte, Christoph Wierling, Hans Lehrach, Michael Lappe and Bodo MH Lange. «The Structural Impact of Cancer-Associated Missense Mutations in Oncogenes and Tumor Suppressors». Molecular Cancer 10, n. 1 (2011): 54. https://doi.org/10.1186/1476-4598-10-54.

Stevanin G., Valérie Hahn, Ebba Lohmann, Naima Bouslam, Michel Gouttard, Caroline Soumphonphakdy, Marie-Laure Welter, et al. «Mutation in the Catalytic Domain of Protein Kinase C γ and Extension of the Phenotype Associated With Spinocerebellar Ataxia Type 14». Archives of Neurology 61, n. 8 (2004). https://doi.org/10.1001/archneur.61.8.1242.

Stokoe David and Frank McCormick. «Activation of C-Raf-1 by Ras and Src through Different Mechanisms: Activation in Vivo and in Vitro». The EMBO Journal 16, n. 9 (1997): 2384–96. https://doi.org/10.1093/emboj/16.9.2384.

Stratton M. R., Peter J. Campbell and P. Andrew Futreal. «The Cancer Genome». Nature 458, n. 7239 (2009): 719–24. https://doi.org/10.1038/nature07943.

Sullivan R.J., Jeffrey R. Infante, Filip Janku, Deborah Jean Lee Wong, Jeffrey A. Sosman, Vicki Keedy, Manish R. Patel, et al. «First-in-Class ERK1/2 Inhibitor Ulixertinib (BVD-523) in Patients with MAPK Mutant Advanced Solid Tumors: Results of a Phase I Dose-Escalation and Expansion Study». *Cancer Discovery* 8, n. 2 (2018): 184–95. https://doi.org/10.1158/2159-8290.CD-17-1119.

Tabor H. K., Neil J. Risch and Richard M. Myers. «Candidate-Gene Approaches for Studying Complex Genetic Traits: Practical Considerations». Nature Reviews Genetics 3, n. 5 (2002): 391–97. https://doi.org/10.1038/nrg796.

Takamiya O., M. Seta, K. Tanaka, e F. Ishida. «Human Factor VII Deficiency Caused by S339C Mutation Located Adjacent to the Specificity Pocket of the Catalytic Domain: Factor VII Deficiency Caused by a Variant S339C». Clinical & Laboratory Haematology 24, n. 4 (2002): 233–38. https://doi.org/10.1046/j.1365-2257.2002.00449.x.

Tanoue T. "Identification of a Docking Groove on ERK and p38 MAP Kinases That Regulates the Specificity of Docking Interactions." The EMBO Journal 20, no. 3 (2001): 466–79. https://doi.org/10.1093/emboj/20.3.466.

Taylor C. A., Kevin W. Cormier, Shannon E. Keenan, Svetlana Earnest, Steve Stippec, Chonlarat Wichaidit, Yu-Chi Juang, et al. "Functional Divergence Caused by Mutations in an Energetic Hotspot in ERK2." Proceedings of the National Academy of Sciences 116, no. 31 (2019): 15514–23. https://doi.org/10.1073/pnas.1905015116.

Taylor Susan S. and Alexandr P. Kornev. "Protein Kinases: Evolution of Dynamic Regulatory Proteins." Trends in Biochemical Sciences 36, no. 2 (2011): 65–77. https://doi.org/10.1016/j.tibs.2010.09.006.

Terrell M.E. and Deborah K. Morrison. «Ras-Mediated Activation of the Raf Family Kinases». Cold Spring Harbor Perspectives in Medicine 9, n. 1 (2019): a033746. https://doi.org/10.1101/cshperspect.a033746.

The International Cancer Genome Consortium. «International Network of Cancer Genome Projects». Nature 464, n. 7291 (2010): 993–98. https://doi.org/10.1038/nature08987.

The International HapMap Consortium. «A Second Generation Human Haplotype Map of over 3.1 Million SNPs». Nature 449, n. 7164 (2007): 851–61. https://doi.org/10.1038/nature06258.

The ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium, «Pan-cancer analysis of whole genomes». Nature volume 578, (2020): 82–93.

Turjanski A.G., J P Vaqué, and J S Gutkind. "MAP Kinases and the Control of Nuclear Events." Oncogene 26, no. 22 (2007): 3240–53. https://doi.org/10.1038/sj.onc.1210415.

Turk Benjamin E. «Understanding and Exploiting Substrate Recognition by Protein Kinases». Current Opinion in Chemical Biology 12, n. 1 (2008): 4–10. https://doi.org/10.1016/j.cbpa.2008.01.018.

Ubersax Jeffrey A. James E. Ferrell Jr. «Mechanisms of Specificity in Protein Phosphorylation». Nature Reviews Molecular Cell Biology 8, n. 7 (2007): 530–41. https://doi.org/10.1038/nrm2203.

Uzun A., C. M. Leslin, A. Abyzov and V. Ilyin. «Structure SNP (StSNP): A Web Server for Mapping and Modeling NsSNPs on Protein Structures with Linkage to Metabolic Pathways». Nucleic Acids Research 35, n. Web Server (2007): W384–92. https://doi.org/10.1093/nar/gkm232.

Vandamme D., Herrero A., Al-Mulla F. and Kolch W. «Regulation of the MAPK Pathway by Raf Kinase Inhibitory Protein». Critical Reviews in

Oncogenesis 19, n. 6 (2014): 405–15. https://doi.org/10.1615/CritRevOncog.2014011922.

Vantaggiato Chiara, Ivan Formentini, Attilio Bondanza, Chiara Bonini, Luigi Naldini, e Riccardo Brambilla. «ERK1 and ERK2 Mitogen-Activated Protein Kinases Affect Ras-Dependent Cell Signaling Differentially». Journal of Biology 5, n. 5 (2006): 14. https://doi.org/10.1186/jbiol38.

Varga A., Jean-Charles Soria, Antoine Hollebecque, Patricia LoRusso, Johanna Bendell, Shih-Min A. Huang, Marie-Claire Wagle, et al. "A First-in-Human Phase I Study to Evaluate the ERK1/2 Inhibitor GDC-0994 in Patients with Advanced Solid Tumors." Clinical Cancer Research 26, no. 6 (2020): 1229–36. https://doi.org/10.1158/1078-0432.CCR-19-2574.

Venselaar H., Franscesca Camilli, Shima Gholizadeh, Marlou Snelleman, Han G Brunner and Gert Vriend. «Status Quo of Annotation of Human Disease Variants». BMC Bioinformatics 14, n. 1 (2013): 352. https://doi.org/10.1186/1471-2105-14-352.

Wang Zhen and John Moult. «SNPs, Protein Structure, and Disease». Human Mutation 17, n. 4 (2001): 263–70. https://doi.org/10.1002/humu.22.

Wang, Zhen, and John Moult. "Three-Dimensional Structural Location and Molecular Functional Effects of Missense SNPs in the T Cell Receptor V? Domain." Proteins: Structure, Function, and Genetics 53, no. 3 (2003): 748–57. https://doi.org/10.1002/prot.10522.

Wold William and Karoly Toth. "Adenovirus Vectors for Gene Therapy, Vaccination and Cancer Gene Therapy." Current Gene Therapy 13, no. 6 (2014): 421–33. https://doi.org/10.2174/1566523213666131125095046.

Xu K., Yan Gao, Li Yang, Yanju Liu, and Chunli Wang. “Magnolin Exhibits Anti-Inflammatory Effects on Chondrocytes via the NF-κB Pathway for Attenuating Anterior Cruciate Ligament Transection-Induced Osteoarthritis.” Connective Tissue Research 62, no. 4 (2021): 475–84. https://doi.org/10.1080/03008207.2020.1778679.

Yamada Y., Yoshiko Banno, Hitoshi Yoshida, Ryosuke Kikuchi, Yukihiro Akao, Takashi Murate, e Yoshinori Nozawa. «Catalytic Inactivation of Human Phospholipase D2 by a Naturally Occurring Gly901Asp Mutation». Archives of Medical Research 37, n. 6 (2006): 696–99. https://doi.org/10.1016/j.arcmed.2006.01.006.

Yao Y., W. Li, J. Wu, U. A. Germann, M. S. S. Su, K. Kuida, e D. M. Boucher. «Extracellular Signal-Regulated Kinase 2 Is Necessary for Mesoderm Differentiation». Proceedings of the National Academy of Sciences 100, n. 22 (2003): 12759–64. https://doi.org/10.1073/pnas.2134254100.

Yates C.M. and Michael J.E. Sternberg. «The Effects of Non-Synonymous Single Nucleotide Polymorphisms (NsSNPs) on Protein–Protein Interactions». Journal of Molecular Biology 425, n. 21 (2013): 3949–63. https://doi.org/10.1016/j.jmb.2013.07.012.

Ye Y., Zhanwen Li, e Adam Godzik. «Modeling and analyzing three-dimensional structures of human disease proteins». In Biocomputing (2006): 439–50. Maui, Hawaii: WORLD SCIENTIFIC, 2005. https://doi.org/10.1142/9789812701626_0040

Yoon Seunghee and Rony Seger. «The Extracellular Signal-Regulated Kinase: Multiple Substrates Regulate Diverse Cellular Functions». Growth Factors 24, n. 1 (2006): 21–44. https://doi.org/10.1080/02699050500284218.

Yuan H.-Y.J.-J. Chiou, W.-H. Tseng, C.-H. Liu, C.-K. Liu, Y.-J. Lin, H.-H. Wang, A. Yao, Y.-T. Chen, e C.-N. Hsu. «FASTSNP: An Always up-to-Date and Extendable Service for SNP Function Analysis and Prioritization». Nucleic Acids Research 34, n. Web Server (2006): W635–41. https://doi.org/10.1093/nar/gkl236.

Yue P., Zhaolong Li, e John Moult. «Loss of Protein Structure Stability as a Major Causative Factor in Monogenic Disease». Journal of Molecular Biology 353, n. 2 (2005): 459–73. https://doi.org/10.1016/j.jmb.2005.08.020.

Yue Peng and John Moult. «Identification and Analysis of Deleterious Human SNPs». Journal of Molecular Biology 356, n. 5 (2006): 1263–74. https://doi.org/10.1016/j.jmb.2005.12.025.

Zhang J., Priscilla L. Yang and Nathanael S. Gray. «Targeting Cancer with Small Molecule Kinase Inhibitors». Nature Reviews Cancer 9, n. 1 (2009): 28–39. https://doi.org/10.1038/nrc2559.

Zhang, Yi, Mohith Manjunath, Jialu Yan, Brittany A. Baur, Shilu Zhang, Sushmita Roy, and Jun S. Song. "The Cancer-Associated Genetic Variant Rs3903072 Modulates Immune Cells in the Tumor Microenvironment." Frontiers in Genetics 10 (2019). https://doi.org/10.3389/fgene.2019.00754.

Zhang Z., Shaolei Teng, Liangjiang Wang, Charles E. Schwartz, e Emil Alexov. «Computational Analysis of Missense Mutations Causing Snyder-Robinson Syndrome». Human Mutation 31, n. 9 (2010): 1043–49. https://doi.org/10.1002/humu.21310.

Zhang Z., Susumu Kobayashi, Alain C. Borczuk, Rom S. Leidner, Thomas LaFramboise, Alan D. Levine, and Balazs Halmos. "Dual Specificity Phosphatase 6 (DUSP6) Is an ETS-Regulated Negative Feedback Mediator of

Oncogenic ERK Signaling in Lung Cancer Cells." Carcinogenesis 31, no. 4 (2010): 577–86. https://doi.org/10.1093/carcin/bgq020.

Zhou B., Zhi-Xin Wang, Yu Zhao, David L. Brautigan, and Zhong-Yin Zhang. "The Specificity of Extracellular Signal-Regulated Kinase 2 Dephosphorylation by Protein Phosphatases." Journal of Biological Chemistry 277, no. 35 (2002): 31818–25. https://doi.org/10.1074/jbc.M203969200.

Zhou B., Channing J. Der and Adrienne D. Cox. «The Role of Wild Type RAS Isoforms in Cancer». Seminars in Cell & Developmental Biology 58 (2016): 60–69. https://doi.org/10.1016/j.semcdb.2016.07.012.

Zhou T., Yuxia Zhang, Antonio Macchiarulo, Zhihong Yang, Marco Cellanetti, Eliecer Coto, Pingyi Xu, Roberto Pellicciari and Li Wang. «Novel Polymorphisms of Nuclear Receptor SHP Associated with Functional and Structural Changes». Journal of Biological Chemistry 285, n. 32 (2010): 24871–81. https://doi.org/10.1074/jbc.M110.133280.

Zuccotto F., Elena Ardini, Elena Casale and Mauro Angiolini. «Through the "Gatekeeper Door": Exploiting the Active Kinase Conformation». Journal of Medicinal Chemistry 53, n. 7 (2010): 2681–94. https://doi.org/10.1021/jm901443h.