UNIVERSITAT AUTÒNOMA DE BARCELONA
Departament de Bioquímica i Biologia Molecular
Facultat de Medicina

TESI DOCTORAL

NEUROKININ-1 RECEPTOR: STRATEGIES OF EXPRESSION, PURIFICATION AND REFOLDING

Memòria presentada per Mikhail Orel per optar al grau de doctor.

Aquest treball ha estat realitzat a la Unitat de Biofísica i el Centre d'Estudis en Biofísica (CEB) del Departament de Bioquímica i Biologia Molecular de la Universitat Autònoma de Barcelona, sota la direcció del Dr. Joan Manyosa Ribatallada.

Vist-i-plau del director de la tesi:

Dr. Joan Manyosa Ribatallada

Bellaterra, Març 2012

Lo más bello que podemos experimentar es el lado misterioso de la vida.
Es el sentimiento profundo que se encuentra en la cuna
del arte y de la ciencia verdadera.

Albert Einstein

INDEX

LIST OF ABBREVIATIONS

Ala	alanine	**Leu**	leucine
Arg	arginine	**Lys**	lysine
Asn	asparagine	**Met**	methionine
Asp	aspartic acid	**Phe**	phenylalanine
Cys	cysteine	**Pro**	proline
Gln	glutamine	**Ser**	serine
Glu	glutamic acid	**Thr**	threonine
Gly	glycine	**Trp**	tryptophan
His	histidine	**Tyr**	tyrosine
Ile	isoleucine	**Val**	valine

aa	amino acid
AEBSF	4-(2-Aminoethyl) benzenesulfonyl fluoride hydrochloride
ALL	acute lymphoblastic leukemia
AMP	adenosine monophosphate
APS	ammonium persulfate
ATR	attenuated total reflection
AU	absorbance units
b2AR	*β*-2 adrenergic receptor
BEVS	baculovirus vector expression system
***β*-Gal**	*β*-galactosidase
bp	base pairs
BSA	bovine serum albumin
cAMP	cyclic adenosine monophosphate
CB1	cannabinoid 1 receptor
CD	circular dichroism
cDNA	complementary DNA
DMEM	Dubelcco's modified eagle medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
EC	endothelial cell
ECL kit	enhanced chemiluminescence kit
E. coli	Escherichia coli
EDTA	ethylenediaminetetraacetic acid
FBS	fetal bovine serum
FTIR	fourier transform infrared spectroscopy
GHS	reduced gluthatione
GLP-1R	glucagon-like peptide-1 receptor
GSSG	oxidized glutathione
GST	glutathione S-transferase
GIP	G protein-coupled receptor Interacting protein
GPCR	G protein-coupled receptor
GuHCl	guanidinium hydrochloride
IBs	inclusion bodies

IMAC	immobilized-metal affinity chromatography
IPTG	isopropyl-*β*-D-thiogalactopyranoside
LB	Luria Bertani medium
LBA	Luria Bertani medium supplemented with ampicillin
MBP	maltose binding protein
MW	molecular weight
NK1R	neurokinin-1 Receptor
NKA	neurokinin A (Substance K)
NMR	nuclear magnetic resonance transmembrane
NPY	neuropeptide Y
OD	optical density
o/n	overnight
PCR	polymerase chain reaction
PEI	polyethylenimine
r-hGH	human growth hormone
RNA	ribonucleic acid
RT	room temperature
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
SP	substance P
TACR	tachykinin receptor
TCEP	tris(2-carboxyethyl)phosphine
TEMED	tetramethylethylenediamine
TM	transmembrane domain
tRNA	transport ribonucleic acid
UV	ultraviolet
VEGF	vascular endothelial growth factor

INTRODUCTION 1

1.1. G protein-coupled receptors

1.1.1. General features of GPCRs and their classification

G protein-coupled receptors (GPCRs) is a large family of membrane proteins involved in the recognition and transduction of signals mediated by different stimuli as diverse as light, Ca^{2+} ions, odorants, amino-acid residues, nucleotides, peptides and proteins (Bockaert and Pin, 1999). A recent and detailed analysis of the human genome reveals over 800 unique GPCRs. Based on sequence similarity within the 7 TM segments, these receptors can be clustered into 5 families: the rhodopsin family (701 members), the adhesion family (24 members), the frizzled/taste family (24 members), the glutamate family (15 members), and the secretin family (15 members). The physiologic function of a large fraction of these 800 GPCRs is fairly unknown (Kobilka, 2007).

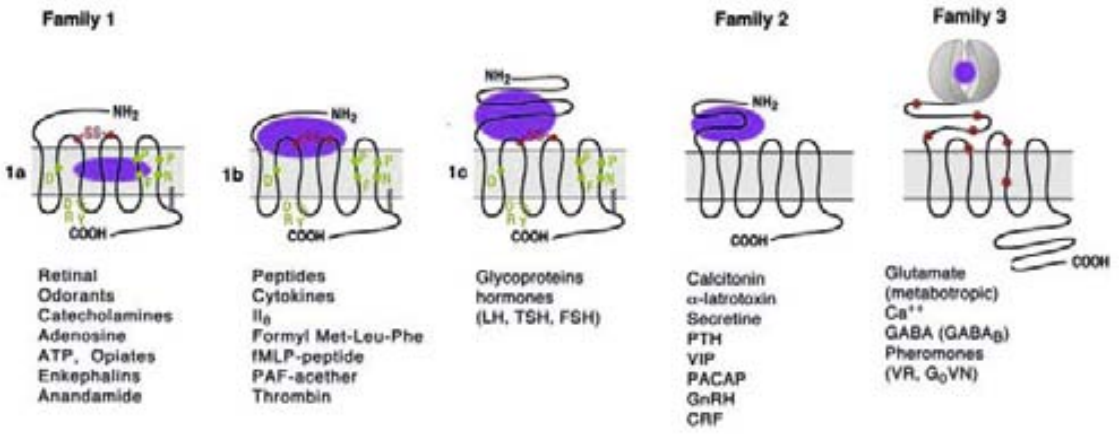


Figure 1.1 Classification of GPCRs (Bockaert and Pin, 1999). Three main GPCRs families are shown.

Another classification of GPCRs is based on the similarity of their morphology and functions (figure 1.1). Receptors from different families show a remarkable example of molecular convergence as they share no sequence similarity. The Family 1 contains most GPCRs including receptors for odorants. Group 1A consists of GPCRs for small ligands and includes rhodopsin and β-adrenergic receptors. The binding site of these two receptors is found within the 7-helix transmembrane core. Group 1B contains receptors for peptides (Neurokinin-1 receptor) whose binding site includes the N-terminal, the extracellular loops and the superior parts of the transmembrane core. Group 1C includes GPCRs for

glycoprotein hormones. Family 2 is very similar to group 1C in its morphology. Their ligands include high molecular weight hormones such as glucagon and secretin. Family 3 contains mGluRs, pheromones and the Ca^{2+} sensing receptors (Bockaert and Pin, 1999).

1.1.2. Structural characteristics of GPCRs

GPCRs are integral membrane proteins that share seven membrane-spanning domains in their structure (figure 1.2). The greatest insight into structure/function relationships of GPCRs can be obtained from comparing *β*2AR and rhodopsin structures (Kobilka and Schertler, 2008).

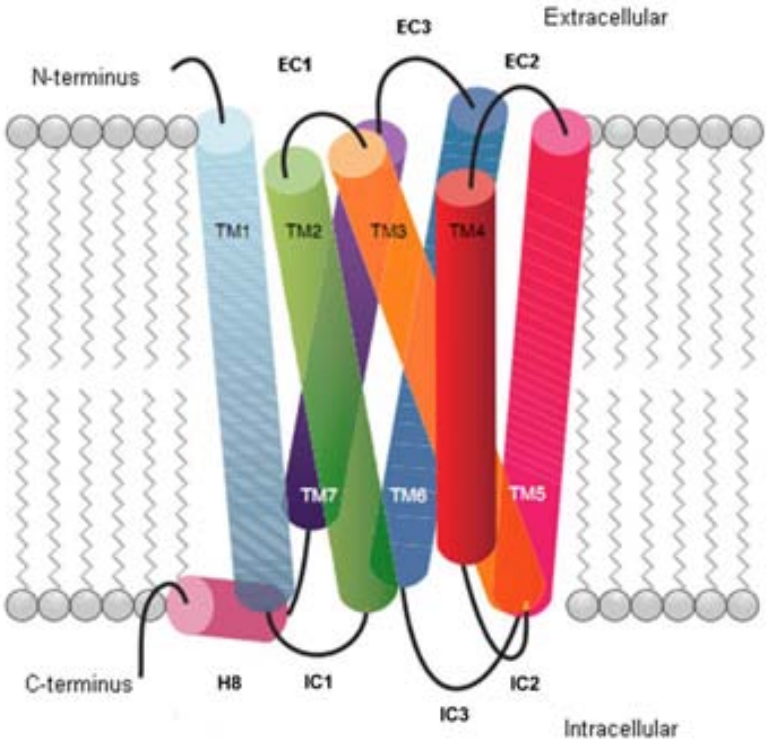


Figure 1.2 General structure of a GPCR embedded in the cell membrane. Lipid bilayer is colored in grey, *TM* – transmembrane helix, *IC* – intracellular loops, *EC* – extracellular loops (Mobarec and Filizola, 2008).

The trans-membrane domain of GPCRs is characterized by several highly conserved residues within the transmembrane helixes and palmytoylated cysteine located near the C-terminus. There is usually a disulfide bond formed between the loop EC2 and the top of transmembrane helix III (Tebben and Schnur, 2011). N-terminus of GPCRs has a great diversity in its amino acid sequence and can

vary in length from relatively short (10–50 amino acids) for monoamine and peptide receptors to much larger (350–600 amino acids) for glycoprotein hormone receptors, and the glutamate family receptors. GPCRs C-terminus sometimes called the “magic tail”, is remarkable for the fact that more than 50 GIPs (GPCR-interacting proteins) have been identified as binding partners of this domain (Bockaert et al., 2003).

All three extracellular loops of GPCRs are extremely important for the correct receptor activation. The loop EC1 is believed to provide structure to the extracellular region of the ligand binding site and enables movement of the transmembrane helices upon ligand binding. The loop EC2 is directly involved in the ligand’s entry, recognition and binding. The loop EC3 is known to be essential for appropriate folding and signaling of the receptor and might be a binding site itself (Peeters et al., 2010). Intracellular loops of GPCRs as, for example, the loop IC3 in thyrotropin-releasing hormone receptor type 1 (Buck et al., 2000), are reported to be responsible for the proper interaction of the receptor with the correspondent G protein.

For some GPCRs the formation of dimers or oligomeric forms as, for example, for β_2-adrenergic (Herbert et al., 1996), M3 muscarinic receptor (Zeng and Wess, 1999) and rhodopsin (Fotiadis et al., 2003 and Liang et al., 2003) was observed. Also, the heterodimeric formation is typical for α_{1a}- and α_{1b}-adrenergic receptors (Stanasila et al., 2003).

1.1.3. Mechanisms of signal transduction

GPCRs are among the essential nodes of communication between the internal and external environments of cells. All of GPCRs present very similar mechanism of activation initiated by binding of a ligand which activates specific heterotrimeric proteins, leading to the modulation of downstream effector proteins (figure 1.3). As an example, in human β_2-adrenergic receptor, the binding of adrenaline and noradrenaline leads to the activation of the stimulatory subunit of the heterotrimeric G protein (Gαs), the stimulation of adenylyl cyclase, the accumulation of cyclic AMP (cAMP), the activation of cAMP-dependent protein kinase A (PKA) and the

phosphorylation of proteins involved in muscle-cell contraction (Rosenbaum, 2009).

It is believed that the conformational changes produced in a receptor upon the binding of its ligand affect the conformation of IC2 and IC3 loops that are known to be key points of G protein recognition and GPCR activation.

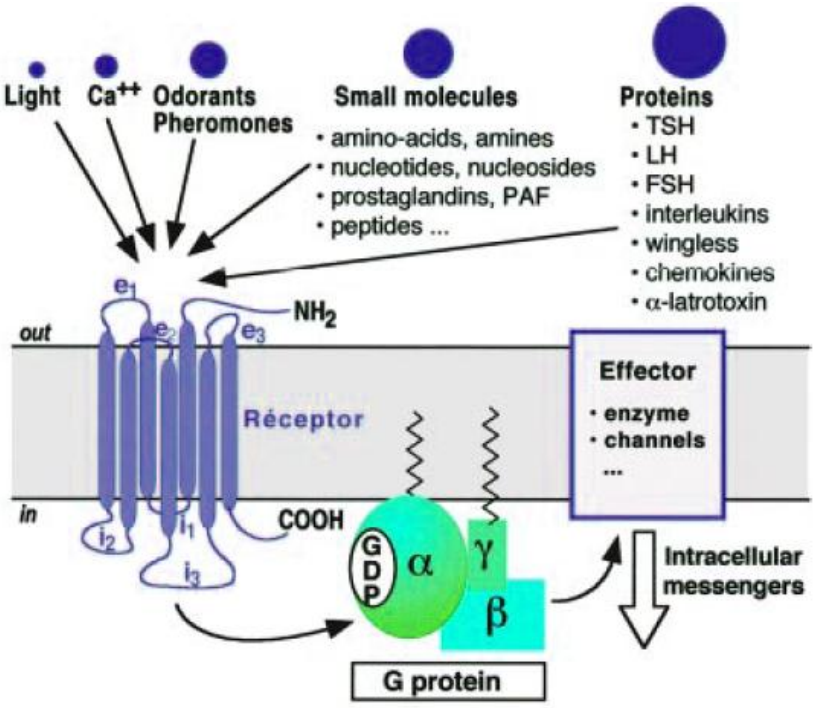


Figure 1.3 General scheme of GPCRs activation pathway upon ligand binding (Bockaert and Pin, 1999).

For instance, in rhodopsin the loop IC2 plays an important role in the receptor activation whereas the IC3 loop is responsible for the interaction and specificity. In β_2-adrenergic receptor both the IC2 and IC3 are involved in the receptor activation. The inactivation of the GPCRs is carried by different kinases able to phosphorylate the receptor as rhodopsin kinase does in case of rhodopsin. The alternative mechanism to phosphorylation is receptor internalization.

1.2. Tachykinin peptides family

Tachykinins constitute the largest peptide family, which play an important regulatory role in vertebrates. Three peptides of this family, substance P, neurokinin A and neurokinin B have an established role as neurotransmitters in mammals (Maggi, 1995). Since Substance P was discovered as a structurally unidentified vasodilatory and contractile compound in 1931, the tremendous advances have been made regarding molecular and functional characterization of tachykinins and their receptors, revealing diverse molecular species of TK peptides (Satake and Kawada, 2006). There are three known mammalian tachykinin receptors termed neurokinin-1 (NK1R), neurokinin-2 (NK2R) and neurokinin-3 (NK3R) receptors. All are members of the 7 transmembrane G-protein coupled receptor family and induce the activation of phospholipase C, producing inositol triphosphate (so called G_q-coupled).

1.2.1. General characteristics, localization and functions of tachykinin peptides

The tachykinin family is characterized by a common C-terminal sequence, Phe-X-Gly-Leu-Met-NH2, where X is either an aromatic or an aliphatic amino acid. Only three tachykinins have been isolated and sequenced from mammalian tissues until now: SP, NKA and NKB (figure 1.4).

Substance P	*Arg Pro Lys Pro Gln Gln **Phe** Phe **Gly Leu Met** – NH_2*
NKA	*His Lys Thr Asp Ser **Phe** Val **Gly Leu Met** – NH_2*
NKB	*Asp Met His Asp Phe **Phe** Val **Gly Leu Met** – NH_2*

Figure 1.4 Primary sequence of three known tachykinin peptides.

NKA is present in two elongated forms, neuropeptide K and neuropeptide-1. Mammalian tachykinins are derived from two preprotachykinin genes: PTT-A and PPT-B. The PPT-A gene encodes the sequences of SP, NKA, and neuropeptide K and neuropeptide-1, whereas the PPT-B gene encodes the sequence of NKB. The RNA precursor from PPT-A is alternatively processed to yield three different mRNAs (Nawa et al., 1983; Kotani et al., 1986). Tachykinins are liberated from their

precursors by the action of specific processing proteases. Typical cleavage points are Lys-Arg, Arg-Arg, and Arg-Lys doublets and the cleavage is carried out by six groups of proteolytic enzymes called convertases (Steiner et al., 1992).

Like all other neurotransmitters, neuronal tachykinins are released from the nerve endings in response to application of physiological and non-physiological stimuli. Concerning release, two points are firmly established. The release of neuropeptides is considered to be “slow” and is probably discrete and long lasting. At the nerve terminals, especially in brain and in the autonomic nervous system, a release of a single transmitter is improbable hence the concept of co-release of different peptides, amines, amino acids, and purines is now generally accepted (Hokfelt et al., 1986). Once released, the tachykinins may be attacked, cleaved, and inactivated by a number of proteolytic enzymes. For example in SP, (seems to be the most vulnerable peptide) degradation three major enzymes seem to display a predominant role: dipeptidyl-amino peptidase, postproline endopeptidase, and cathepsin D (Regoli et al., 1994).

All of the non-neuronal localizations of the tachykinins concern, mainly SP, and the occurrence of the peptide were established only by immunohistochemistry, using selective SP-antisera. SP occurring in populations of cells of the mammalian intestine may act either as paracrine hormone or as a true hormone after release into the blood stream. It is possible that SP found in mammalian blood originates prevalently from the SP-containing cells of the intestinal mucosa.

Nervous tissue represents by far the most important localization of the tachykinins in invertebrates and vertebrates. A more or less dense network of tachykininergic fibers, which release their content upon adequate stimulation, permeates all vertebrate tissues close upon a very rich population of different receptors located on the membrane of neuronal and non-neuronal cells. In certain cerebral areas, the concentration of tachykinins may be on the order of nanomoles (Hunter et al., 1985).

The most quantitatively important source of tachykinins in the gut is the enteric nervous system, which has its cells in the wall of the intestine and supplies all gastrointestinal effector systems. The mammalian gastrointestinal tract contains both SP and NKA and various extended forms of these tachykinins.

It has been demonstrated the presence of SP and NKA in the respiratory tract of various mammalian fibers. They were found in the trachea and bronchi, in the bronchial tree where most of the SP-positive fibers are of vagal origin and in lungs, the fibers are both of vagal and thoracic spinal origin (Manzini et al., 1989). The distribution of SP and NKA has been extensively studied in renal pelvis and ureter and especially in the urinary bladder of several species (Sharkey et al., 1983; Gibbins et al., 1985). SP-like immunoreactivity is also present in various skin areas, suggesting that SP is present mainly in primary afferent C-fibers (Holzer, 1991). Tachykinin-containing nerves are present in lymphoid organs, such as thymus, spleen, lymph nodes, and lymphoid aggregates in the lung and nasal mucosa.

1.2.2. Substance P. General characteristics and functions

Substance P is an undecapeptide which functions as a neurotransmitter and as a neuromodulator. It was originally discovered in 1931 by Ulf von Euler and John Gaddum in a tissue extract that caused intestinal contraction *in vitro*. The structural determination of the mammalian tachykinin SP by NMR, CD and FTIR have turned out to be inconclusive, although strong arguments for an extended polyproline II structure under defined conditions have been raised (Foroutan et al., 2011). This knowledge is now being extensively used in the development of non-peptide NK1 receptor antagonists that essentially block the pharmacological effects of SP (Datar et al., 2004).

The affinity of fragments of the tachykinin substance P for the cloned rat NK1, NK2, and NK3 receptors suggest that addition of one to three amino-terminal residues to the original sequence results in the optimization of its interaction within the binding pocket of the NK1 receptor. The addition of Pro-Gln-Gln to the C-terminus leads to the increased affinity for the NK1 receptor, either by providing additional binding interactions or by modifying the conformation of the carboxyl-terminal sequence. However the addition of the same sequence to SP is unfavorable for NK2 and NK3 receptor binding (Cascieri et al., 1992).

The biological activity of SP seems to be mediated by the C-terminus of the peptide. The N-terminal fragment (1-7 aa) of SP inhibits nociceptive, aggressive

and grooming behaviors and stimulates investigative motor behavior while C-terminal hexapeptide fragment analog pyroglutamyl-SP (7–11) exerts opposite effects. Some of the behavioral effects produced by exogenous SP exhibit a strikingly different structure-activity relationship.

Substance P is known to be an important element in pain perception and is tightly related to the transmission of pain information into the central nervous system. SP is also associated with the regulation of mood disorders, stress, anxiety, neurogenesis, reinforcement, respiratory rhythm, neurotoxicity, nausea, pain and nociception (Ebner and Singewald, 2006; Huston et al., 1993; Park et al., 2007; Hesketh, 2001; Zubrzycka and Janecka, 2000).

Substance P is involved in neurogenic inflammation – a local inflammatory response to certain types of infection or injury. The regulatory function of SP also involves the regulation of its high-affinity NK1 receptor. Substance P receptor antagonists may have important therapeutic applications in the treatment of a variety of stress-related illnesses, in addition to their potential as analgesics. The particular features of SP, namely the high activity of C-terminus partial sequences in some pharmacological tests were used to develop one hexa- and several octapeptide antagonists. Weak antagonists were obtained with a single modification, namely the replacement of Leu^{10} with Trp on the other hand, [Pro^{4}, $Trp^{7,9,10}$, Phe^{11}]-SP (4-11) was found to be the most potent antagonist of the SP. The data obtained with the octa- and undecapeptide antagonists of SP have been used for identification and characterization of SP receptors in various smooth muscles. It appears that SP and related peptides may exert their numerous pharmacological effects by activating more than one receptor type (Regoli et al., 1984).

1.3. Tachykinin receptors. General characteristics, structure and their functions

The multiple peptides that play a role of neurotransmitters or neuromodulators act mainly by activating three primary types of receptors, NK1, NK2, and NK3 which belong to the family of G protein-coupled receptors (table 1.1). The pharmacology and tissue locations of these receptor sites and their involvement in certain biologic responses are still discussed. Second messenger system established to be activated by tachykinin receptor stimulation includes the hydrolysis of inositol containing phospholipids by a phospholipase C mechanism (Krause et al., 1992).

Table 1.1 Tachykinin receptors types and their ligands.

Receptor name	Gene	Number of aa	Ligands affinity
Neurokinin-1 (NK1R) (Substance P receptor)	TACR1	407	SP > NKA > NKB
Neurokinin-2 (NK2R) (Substance K receptor)	TACR2	390	NKA > NKB > SP
Neurokinin-3 (NK3R) (Neuromedin K receptor)	TACR3	452	NKB > NKA > SP

The similarities and differences in biological properties of the three receptors result from their primary structure sequences. The rat substance P, substance K, and neuromedin K receptors consist of 407, 390, and 452 amino acid residues, respectively, while the bovine substance K receptor is composed of 384 amino acid residues lacking six amino acid residues at the C-terminus (Masu et al., 1987). The expression of the three tachykinin receptor genes may be regulated differently depending on the external stimuli or different development stages. The different distribution of the three tachykinin receptor mRNAs plays and important role in better understanding of the functions of the tachykinin peptides (Nakanishi, 1991).

Regarding the structural characteristics of tachykinin receptors, all of them have seven hydrophobic segments, each consisting of about 20-25 uncharged amino acid residues (Dixon et al., 1988; Bonner, 1989). The transmembrane core is highly conserved among the three tachykinin receptors (Shigemoto et al, 1990; Nakanishi

et al 1990), and the homology in these segments is around (54-66%). The potential N-glycosylation sites are located at the N-end of the receptor, while, many serine and threonine residues which are possible phosphorylation sites are found at the C-terminus. TM2 segments of both the NK2R and NK3R receptors contain an aspartic acid, as observed in other G protein-coupled receptors which is replaced with a glutamic acid in the substance P receptor. All three receptors contain one histidine residue each in both TM5 and TM6 segments, and the presence of these two histidine residues is characteristic of the tachykinin receptors (Nakanishi et al., 1990). The tachykinin receptors are thought to express their functions by activating a phosphatidylinositol-calcium second message system.

1.3.1. Neurokinin-1 receptor (NK1R)

1.3.1.1. General and structural characteristics of NK1R

The tachykinin receptor 1 (TACR1) also known as Neurokinin-1 receptor (NK1R) is a G protein-coupled receptor found in the central nervous system and peripheral nervous system and is encoded by *TACR1* gene. Substance P is known as a major endogenous ligand for this receptor, although it has some affinity for other tachykinins. The protein NK1R is responsible for many physiological processes, which include pain transmission, exocrine and endocrine secretion, vasodilatation, modulation of cell proliferation, regulation of the immune and inflammatory responses (Tuluc et al., 2009).

The genomic DNA isolated for NK1R extends over 60 kb. The gene consists of 5 exons interrupted by 4 introns located within the protein-coding regions. (Takahashi et al., 1992). Comparison of different species of NK1R as, for example, rat and human sequences demonstrates about 94.5% identity.

Neurokinin-1 receptor has 7 *α*-helical transmembrane domains like all other G protein-coupled receptors (figure 1.5). The human NK1R has 2 consensus N-linked glycosylation sites in the amino terminal domain, and a potential palmitoylation site at Cys^{323}. Multiple potential serine and threonine phosphorylation sites exist in the IC3 loop and C-terminus domain (Takeda et al., 1991).

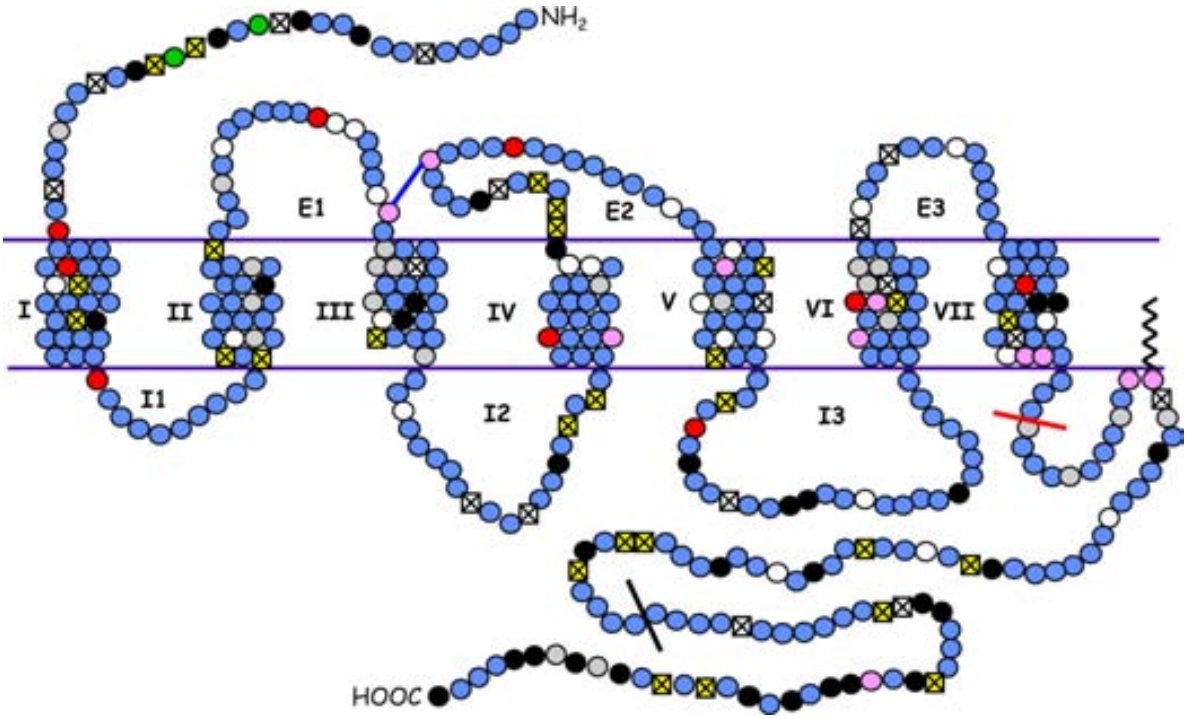


Figure 1.5 Residue-based diagram of Neurokinin-1 receptor. I1, I2 and I3 – cytoplasmic loops; E1, E2 and E3 – extracellular loops; I, II, III, VI, V, VI, VII – transmembrane helices by their number. Fluorescent amino acids: – phenylalanine, – tyrosine, – tryptophan; – serine which can undergo phosphorylation; – threonine; – proline; – aspartic; – cysteine, – disulphide bonds between two cysteines. *Black slash bar* – hNK1R-366 truncated form*; red slash bar* – hNK1R-311 truncated form.

Since tachykinin peptides are relatively large molecules, the peptide binding domain of NK1R may involve not only the transmembrane core, but some additional residues. The substitution of the residues 1-13 of the rat NK1R has a small effect on peptide binding, however, the similar substitution extended to residues 1-27, leads to the loss of high-affinity binding of SP. This decrease in binding affinity could result from the removal of residues critical for peptide binding, or from the introduction of the non-NK1R residues at positions homologous to residues 14-27, which are required for the high-affinity binding of peptides. It was also demonstrated that the residues 21-29 are required for the high-affinity interaction between the NK1R and the peptides.

The studies on the ligand binding site of the NK1R show that the peptide binding site probably involves at least nine residues in TM2, TM5 and extracellular domains (Asn^{23}, Gln^{24}, Phe^{25}, Asn^{85}, Asn^{89}, Tyr^{92}, Asn^{96}, His^{108}, and Tyr^{287}), while the binding site for small molecule antagonists includes at least three residues in

the transmembrane domains (His^{197}, His^{265}, and Tyr^{287}) (Fong et al., 1992). The nature of the interactions between the peptides and Asn^{85}, Asn^{89}, or Asn^{96} of the NK1R is likely to be hydrogen-bonding interactions while Tyr^{92} appears to be involved in an aromatic interaction with the peptides because the Y92F mutant receptor is functionally identical to the wild-type receptor. Analysis of several SP analogs suggests that the C-terminus amide group of SP may interact with Asn^{85} in the second transmembrane domain, whereas Glu^{78} and Tyr^{205} in the second and fifth transmembrane domains are critical for receptor activation. In contrast, TM5, TM6 and TM7 appear to make direct contact with non-peptide antagonists and are likely to be near the bound peptides (Huang et al., 1994).

1.3.1.2. Isoforms of Neurokinin-1 receptor

Two naturally occurring variants of the NK1R mediate the effects of SP: a full-length receptor and a truncated (tail-less) form that lacks 96 amino acid residues at the C-terminus (NK1R-311 form) (figure 1.6). Most of the functions of NK1R are assumed to be mediated by the full-length form of the receptor, however, there are some studies aimed to elucidate the functional responses mediated by the truncated NK1R.

An artificially truncated NK1R lacking the last 82 amino acid residues in the C-terminus mediates calcium mobilization in response to stimulation with agonists, suggesting that none of the residues between Phe^{325} and Ser^{407} is essential for the coupling of full length NK1R with downstream signals via G_q/11 proteins (Bohm et al., 1997).

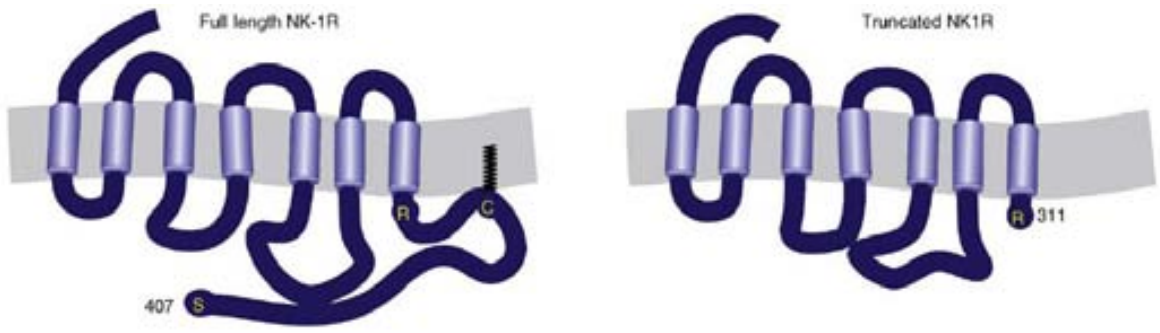


Figure 1.6 Two naturally occurring isoforms of NK1 receptor: the full-length form and 311-407 truncated form are shown (from Tuluc et al, 2009).

It is believed that the presence of the C-terminus domain has an impact on the folding of NK1R, and the lack of the C-terminus confers to the truncated receptor distinct pharmacologic properties, despite the identity between the primary structure of the truncated receptor with the first 311 amino acids in the sequence of the full-length receptor. The truncated form of NK1R most likely mediates prolonged cellular responsiveness after stimulation, and is resistant to homologous desensitization in comparison to the full length NK1R isoform because the missing C-terminus domain is essential for receptor desensitization, internalization and recycling (Tuluc et al., 2009).

1.3.1.3. Functional aspects of Neurokinin-1 receptor

Neurokinin-1 receptor is considered a promising drug target, particularly with regards to potential analgesics and anti-depressants; also it was identified as a candidate in the etiology of bipolar disorder. Furthermore NK1R antagonists may be used for cancer and alcoholism treatment. The NK1R receptor is tightly related to cancer due to its overexpression in tumors (cancer cells express more NK1R receptors than normal cells). NK1R receptor antagonists specifically inhibit tumor cell proliferation, angiogenesis and the migration of the tumor cells (Muñoz and Cavenas, 2010). In acute lymphoblastic leukemia (ALL), SP is expressed in human blast cells. Using a knockdown method, it was demonstrated that NK1R is involved in the viability of tumor cells and some of its antagonists are able to inhibit ALL cell growth, which suggests that NK1R antagonists could be considered as new antitumor drugs for the treatment of human ALL (Muñoz et al., 2012). NK1R and SP are believed to be related to the immune system and selected infections. Substance P binding to NK1R, leads to the activation of nuclear factor-kappa B and proinflammatory cytokines, which is important in bacterial, viral, fungal, and parasitic diseases, as well as in immune system function (Douglas and Leeman, 2010).

1.3.1.4. C-terminus domain of Neurokinin-1 receptor

The NK1 receptor C-terminus domain seems to be essential for the coupling to corresponding G protein (Fong et al., 1992; Lai et al., 2006) and *β*-arrestin, (Schmidlin et al., 2003), and is essential for receptor desensitization, internalization and recycling. As it was already mentioned above, an artificially truncated NK1R lacking the last 82 amino acid residues in the C-terminus can still mediate calcium mobilization in response to stimulation with agonists, suggesting that none of the residues between Phe^{325} and Ser^{407} is essential for the coupling of full-length NK1R with downstream signals via Gq/11 proteins (Bohm et al., 1997). Wild-type and truncated forms of the receptor also showed to be indistinguishable in agonist binding affinity and SP-induced accumulation of inositol phosphates (Raddatz et al., 1995). Nevertheless, on the other side, the truncated form of NK1R most likely mediates prolonged cellular responsiveness after stimulation, and is resistant to homologous desensitization in comparison to the full-length NK1R isoform (Tuluc et al., 2009). Despite of the wide range of studies made on NK1R C-terminus role in the processes of the receptor desensitization, phosphorylation and endocytosis, very little information is known about its structure. Those few intents of modeling human Neurokinin-1 receptor structure base on crystal structure of bovine rhodopsin suggest that the NK1R C-terminus has a small initial antiparallel *β*-sheet, spanning segments Thr^{339} to Arg^{340} and Gly^{346} to Ser^{347}. The rest of the C-terminus residues are likely to be fold into a globular structure (Khedkar et al., 2005).

1.4. Obtaining GPCRs for structural studies

G protein-coupled receptors (GPCRs) are targets for about 60–70% of all drugs that are implemented nowadays for treatment of different diseases. Traditionally, the drug discovery process involves the screening of chemical compounds to identify novel and more-efficient drug molecules. However the structure-based drug design can provide a targeted approach in drug, it is severely limited by our knowledge of high resolution structures of GPCRs which arises from the extreme difficulty of their expression, purification and crystallization (Lundstrom, 2006).

In order to obtain sufficient amounts of GPCR for biophysical studies it is crucial to set up an overexpression system for each receptor of interest. Many attempts have been made to improve functional membrane protein expression levels by screening various combinations of factors and conditions in the expression systems, such as host organism, plasmid vector, fusion adducts, codon usage, and expression induction conditions (Sarramegna, 2003). However, even if all of these variables could be globally optimized, they still may not yield enough functional material because of inherent limitations in the protein sequence itself. High resolution structural techniques, such as solid-state NMR and crystallography, require milligram amounts of purified receptor in an active ligand-binding form, and the challenges associated with the overexpression of GPCRs, any optimization of expression levels is clearly advantageous (Sarkar et al., 2008). The next two sections of this dissertate will be dedicated to the review of expression systems implemented for the expression of GPCRs for structural and functional studies.

1.4.1. Mammalian cells as a source of recombinant GPCRs

GPCRs expression in mammalian cells is usually used when it is not necessary to produce big amount of the receptor. A huge number of GPCRs have been expressed in mammalian cells to study the key residues involved in ligand binding as well as possible structure of the active site of the receptor by mutagenesis experiments (Schwartz, 1994).

GPCRs produced in mammalian cells rarely carry any additional tags for a subsequent purification since they are most often dedicated to functional studies (internalization, desensitization or dimerization of the receptor) and not to overproduction. As it can be easily deduced from the table 1.2, different GPCRs may be expressed in different mammalian cell types. COS, HEK and CHO cells lines are of the most common to carry out the expression. The expression level is highly variable from one GPCR to another. The highest production level reported is that of the human β_2-adrenergic receptor, which reached 200 pmol/mg of membrane protein (Lohse, 1992) and rhodopsin with 10 mg/l in culture of HEK-293 cells (Saramegna et al., 2003).

Table 1.2 GPCRs, which were reported to be expressed in mammalian cells.

Receptor	Origin	Cell type	Functionality	References
β_2-Adrenergic	Human	CHO	+	Lohse, 1992
Adenosine A_1	Human	HEK-293 CHO	+ +	Gao et al., 1999 Townsend-Nicholson and Shine, 1992
Calcitonin	Human	MEL	+	Needham et al., 1995
Glutamate $mGlu_{2, 3}$	Rat	CHO/SFV	+	Malherbe et al., 1999
Histamine-H2	Rat	COS-7/SFV	+	Hoffmann et al., 2001
$M_{1,4}$-Muscarinic	Mouse	JEG-3	+	Migeon and Nathanson, 1994
$M_{1,2,3,4}$-Muscarinic	Human	HEK-293	+	Peralta et al., 1988
Neurokinin-1	Human Rat	BHK/SFV COS	+ +	Lundstrom et al., 1994 Ingi et al., 1991
Neurokinin-2	Human	CHO	+	Turcatti et al., 1993
Neuromedin K	Rat	COS	+	Ingi et al., 1991
Neuropeptide Y1	Human	HeLa/VV	+	Walker et al., 1993
Rhodopsin	Bovine	COS-1 HEK-293S	+ +	Oprian et al., 1987 Reeves et al., 2002
Serotonin $5HT_{1E}$	Human	HEK-293	+	McAllister et al., 1992
Serotonin $5HT\text{-}dro_{2A}$	*Drosophila*	COS-7	+	Saudou et al., 1992
Substance K	Human	NIH-3T3	+	Kris et al., 1991

1.4.2. GPCRs expression in *Escherichia coli*

E. coli expression system is a low cost and rapid method for protein expression and can be applied for production of membrane proteins and GPCRs in particular (table1.3). Its main advantages are homogeneity of the recombinant proteins due to the lack of posttranslational modifications, short generation time, high expression rates and short times of induction (Tate and Grisshammer, 1996). In spite all mentioned above, *E. coli* expression system is not the best one for carrying out functional studies of the receptor *in vivo*, since the bacterial cells do not contain endogenous G protein. Moreover, the lack of endogenous G protein can give rise to the absence of high-affinity binding sites for agonists (Stanasila et al., 2000).

In some cases and, for membrane proteins in particular, the reductive environment of the bacterial cytoplasm can slow down the production of functional form of the receptor, especially when the correct folding of the receptor in an active conformation requires formation of disulfide bridges. Sometimes this problem can

be solved by fusing the receptor of interest with *E. coli* periplasmic membrane proteins such as the maltose binding protein (MBP) (Abrahmsen et al., 1986).

Table 1.3 Some GPCRs that were successfully expressed in *E. coli.*

Receptor	Origin	Phusion protein	Receptor localization	Functionality	References
β_1-Adrenergic	Human		Membrane	+	Breyer et al., 1990
		β-Gal	Membrane	+	Marullo et al., 1988
		LamB	Membrane	+	Chapot et al., 1990
		MBP	Membrane	+	Hampe et al., 2000
β_2-Adrenergic	Human	LamB	Membrane	+	Chapot et al., 1990
***μ*-Opioid receptor**	Human		Membrane	+	Stanasila et al., 2000
Adenosine A_1	Human	PhoA	Membrane	+	Lacatena et al., 1994
		MBP		+	
Adenosine A2a	Human	MBP	Membrane	+	Jockers et al., 1994
Endothelin ETB	Human	None	Membrane	+	Haendler et al., 1993
M2 mAch	Human	MBP	Membrane	+	Furukawa and Haga, 2000
Neuropeptide Y_1	Human	MPB	Membrane	+	Munch et al., 1995
Neurotensin	Rat	MBP	Membrane	+	Hill and Sillence, 1997
Neurokinin-1	Human	None	Inclusion bodies	–	Bane et al., 2007
Olfactive receptor OR5	Human	GST	Inclusion bodies	–	Kiefer et al., 1996
Serotonin $5HT_{1A}$	Human	MBP	Membrane	+	Bertin et al., 1992
TSH	Human	None	ND	ND	Busuttil et al.2001
Vasopressin V2	Human	PhoA/-gal	Membrane	–	Schulein et al., 1996

The lipid composition of *E. coli* membranes is very different from that of eukaryotic cells, and this can greatly affect the binding properties of recombinant receptors. In some cases it is important for the functionality of certain receptors such as for example, oxytocin (Gimpl et al., 1995), transferrin (Nunez and Glass, 1982) and human *μ*-opioid receptors (Hasegawa, 1987) that are strongly dependent on the lipid environment.

The use of alternative *E. coli* expression strains revealed that BL21(DE3) cells yielded an approximate 2.5-fold increase of receptor expression when compared with DH5*α*, which have been traditionally used for its expression. The adapted strains C41(DE3) and C43(DE3), derived from BL21(DE3), have allowed the

expression of many toxic proteins, particularly membrane proteins. This may be, at least in part, due to the formation of intracellular membranes during protein induction (reviewed in Attrill et al., 2009).

The over-expression of GPCRs in *E. coli* is often to be appeared toxic for the host strain and results in either poor expression levels or inclusion bodies formation. The inclusion bodies mainly consist of the protein of interest and are easy to extract and purify. Moreover, the accumulation of the overexpressed protein in aggregated form in the cytoplasm protects it from proteolytic degradation by proteases. Despite all the mammalian membrane receptors expressed in form of inclusion bodies are usually inactive, functional refolding after purification is required (Kiefer et al., 1996).

1.5. Refolding of proteins expressed in *E. coli*

1.5.1. General considerations about protein folding

Protein folding is the physical process by which a newly synthesized polypeptide acquires its characteristic and functional three-dimensional structure, known as a native state. The resulting three-dimensional structure is determined by the amino acid sequence. Folded proteins usually have a hydrophobic core in which side chain packing stabilizes the folded state, and charged or polar side chains occupy the solvent-exposed surface where they interact with surrounding water.

Minimizing the number of hydrophobic side-chains exposed to water is an important driving force behind the folding process (Pace et al., 1996). The correct three-dimensional structure is essential to function; however, some parts of functional proteins may remain unfolded. Failure to fold into the intended shape usually produces inactive misfolded proteins with altered properties, which are known to be responsible for several neurodegenerative and other diseases (Selkoe, 2003).

Membrane proteins are inserted into the membrane with the help of a protein complex called the translocon, which forms a tunnel that allows translocation of regions destined for the *trans* compartment as well as lateral release of hydrophobic segments into the bilayer (Chin et al., 2002). The two-stage model of

Popot and Engelman postulates that in a first stage, individual transmembrane segment stretches form stable *α*-helices in the membrane when inserted. In the second stage they associate to form tertiary contacts at the helix interfaces, resulting in formation of the native three-dimensional structure (Popot and Engelman, 2000).

1.5.2. Inclusion bodies: protein aggregates

Until now very few X-ray structures of classical GPCRs have been resolved, including the β_1- and β_2- adrenergic receptors and the A2A adenosine receptor. This lag in obtaining GPCR structures is due to several tedious steps, required before beginning the first crystallization experiments: protein expression, detergent solubilization, purification, and stabilization (Michalke et al., 2010). Upon overexpression of recombinant proteins, the formation of IBs can be observed in several host systems, for example, prokaryotes, yeast or higher eukaryotes as a consequence of high expression rates. The IBs are very dense particles of aggregated protein, which can be visualized by light or electron microscopy (figure 1.7).

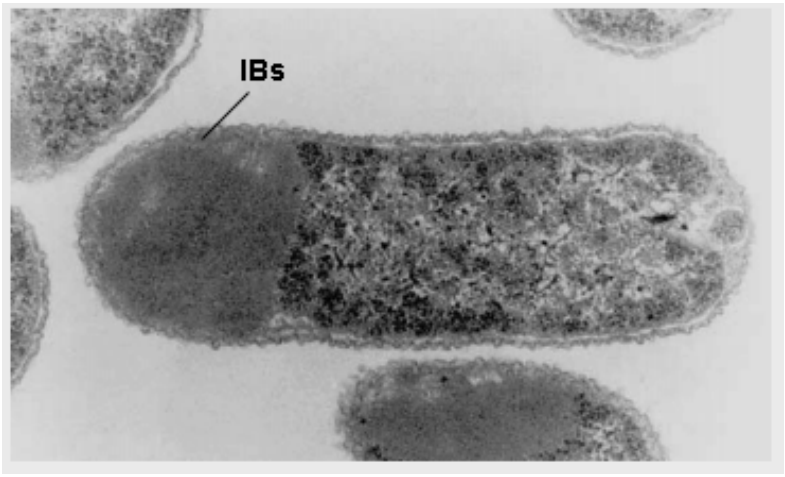


Figure 1.7 Electron micrograph of negatively stained *E. coli* cells displaying inclusion bodies (from Fahnert et al., 2004).

In some cases the protein expressed in IBs can constitute 50% or more of the total cell protein. Major contaminants of IBs material after preparation are outer membrane protein that can be removed by extensive washing with detergents.

There is no direct correlation between the propensity of IBs formation of a certain protein and its intrinsic properties, such as molecular weight, hydrophobicity, folding pathways, etc. Only in case of disulfide bonded proteins the IBs formation can be anticipated. When the protein is produced in the bacterial cytosol, the formation of disulfide bonds usually does not occur in this reducing cellular compartment which results in improper folding and aggregation (reviewed in Lilie et al., 1998).

1.5.3. *In vitro* protein refolding

After successful solubilization of the inclusion bodies in strong detergents or chaotropic agents, the protein becomes partially or fully denatured and cannot carry its functions properly. Refolding process is needed to restore native tertiary structure of a protein along with its native functional state. *In vitro* protein refolding usually consists of the following steps:

1. Growing a cell culture
2. Induction of protein expression
3. Cell lysis and inclusion bodies isolation
4. Inclusion bodies solubilization
5. Screening of suitable refolding conditions
6. Protein refolding
7. Reoxydation to form disulphide bonds
8. Protein purification and/or concentration

1.5.3.1. Refolding by dilution

Protein refolding is initiated by a reduction of denaturant concentration and, by altering the redox environment to enable disulphide bond formation. The simplest and most widely used method for reducing denaturant concentration is by dilution into an appropriate refolding buffer. Two processes occur: first-order refolding and higher order aggregation (Kiefhaber et al., 1991). Refolding at dilute protein concentrations (<10 μg ml^{-1}) can help to minimize the aggregation, however, such low concentrations are impractical for preparative work, so yield is maximized by

altering the mode of denaturant dilution or by using folding enhancers. As no universal refolding buffer can be identified, it is necessary to screen a limited set of conditions (table 1.5) for each protein (Yasuda et al., 1998).

Table 1.5 Some of the most common additives used in refolding buffers (from Middelberg, 2002)

Additive	Typical concentration
Urea	2 M
GuHCl	1 M
Arginine	0.5 M
Glycerol	0.4 M
Sucrose	0.4 M
Lauryl maltoside	0.3 mM
Polyethylene glycol (PEG-3550)	0.05% w/v
Tris buffer	0.5 M
Triton X-100	10 mM
Acetamide	2 M
***n*-hexanol**	2 mM
Salts (NaCl, Na_2SO_4, K_2SO_4)	0.5 M

1.5.3.2. On-column refolding

Refolding using packed columns is an attractive technique because it is easily automated using commercially available preparative chromatography systems. There are three basic approaches:

- immobilization of the denatured protein onto a matrix and subsequent denaturant dilution to promote refolding;
- denaturant dilution using size exclusion chromatography (SEC);
- immobilization of folding protein onto chromatographic supports

The immobilization of denatured protein onto a matrix can be achieved through non-specific interactions or through specific affinity interactions. The aim is to isolate individual protein molecules spatially, thus inhibiting aggregation. Success of the refolding is very protein specific because protein–matrix interactions can prevent refolding. Significant empirical optimization is required for each protein. Nickel-chelating chromatography has also been used for affinity immobilization and

refolding of membrane proteins produced as inclusion bodies in *E. coli* (Rogl et al., 1998) and also for the oxidative refolding of mammalian prion proteins (Zahn et al., 1997).

1.5.4. Successfully expressed and refolded GPCRs produced in *E. coli*

Milligram quantities of neuropeptide Y_2 receptor were produced using high density fermentation with a yield of over 20 mg/L of *E. coli* culture. Extensive studies were carried out on *in vitro* refolding and stabilization of the isolated receptor in detergent solution. The specific binding of the ligand of neuropeptide Y (NPY) to the recombinant Y_2 receptors in micellar form was shown by several radioligand affinity assays (Schmidt et al., 2009).

High-level production in *E. coli* of 4 human GPCRs, chemokine receptors (hCRs) CCR5, CCR3, CXCR4 and CX3CR1 that are directly involved in HIV-1 infection, asthma and cancer metastasis has been reported (Ren et al., 2009).

Another GPCR successfully expressed in *E. coli* is a Vascular endothelial growth factor (VEGF)-A. It interacts with the receptor tyrosine kinases VEGF-R1 and R2, and plays an important role in endothelial cells (EC) function and blood vessel development. To facilitate further characterization of VEGF-B, a protocol for expression and purification of refolded recombinant protein from *E. coli* inclusion bodies has been developed. The resulting purified disulfide-linked homodimer was demonstrated to bind to VEGF-R1 and to compete with VEGF-A for binding to this receptor (Scrofani et al., 2000).

A framework was developed to produce functional human full-length glucagon-like peptide-1 receptor (GLP-1R) by recombinant expression in *E. coli*. Activation of GLP-1R upon ligand binding leads to the release of insulin from pancreatic cells. This strictly glucose-dependent process renders the receptor and its ligands useful in the treatment of type II diabetes mellitus. The ligand binding of the renatured protein was quantified by fluorescence quenching and surface plasmon resonance spectroscopy (SPR) (Schroder-Tittmann et al., 2010).

Purification of inclusion bodies followed by solubilization and refolding was applied for the recovery of Human growth hormone (r-hGH) from *E. coli.* The

protein folds into a four-helix bundle structure with two disulfide bridges. Human growth hormone was expressed as IBs without any additional tag in *E. coli* using the T5 promoter and was used for the recovery of bioactive protein using a novel solubilization procedure (Patra et al., 2000).

OBJECTIVES 2

Neurokinin-1 receptor (NK1R) is a G protein-coupled receptor found in the central and peripheral nervous system of vertebrates and is responsible for many physiological processes. Along with its antagonists, it is considered a promising drug target for many disorders as alcoholism, bipolar disorder, immune system diseases and selected infections. Neurokinin-1 receptor has two naturally occurring forms that differ in the length of the carboxyl terminus: a full-length receptor consisting of 407 amino acids and a truncated 311-residues form, lacking the C-terminus.

High resolution structural techniques, such as solid-state NMR and crystallography, require milligram amounts of purified receptor in an active ligand-binding form, and the challenge associated with the overexpression of GPCRs along with optimization of expression levels is clearly advantageous. The *E. coli* expression system is a low cost and rapid method for GPCRs expression. The first part of this work comprises the obtainment and study of hNK1R-366 and hNK1R-311 truncated forms of hNK1 receptor expressed in *E. coli* and solubilized from inclusion bodies. Radioligand binding assay will be will be implemented as an evidence of proper SP binding to the refolded receptor. We also aim to study the refolded receptor applying various spectroscopic techniques useful for structural characterization of peptides and proteins.

The second part of this study is centered on hNK1R C-terminus expression, purification and characterization to elucidate its structural properties. As it was already mentioned in the introductory section, the C-terminus domain seems to be important for the coupling to corresponding G protein and *β*-arrestin, and is essential for receptor desensitization, internalization and recycling. Circular dichroism along with FTIR spectroscopy are powerful techniques that have been used to study the secondary structure of the peptide. The presence of several tyrosine residues in the primary sequence of hNK1 C-terminus makes possible the use of UV and fluorescence spectroscopy techniques that can be applied for studying the intrinsic conformational changes in the tertiary structure in order to deepen the knowledge of described functional roles of this domain.

The main objectives of the present work are:

1. To set up a system for hNK1R expression and purification in *E. coli* in order to obtain human Neurokinin-1 receptor (hNK1R-366 and hNK1R-311 truncated forms). The suitable solubilization and refolding conditions for expressed hNK1R should be found. The purified and refolded receptor will be characterized applying various spectroscopic techniques useful for structural characterization of peptides and proteins and the radioligand binding assay will be implemented as an evidence of proper SP binding to the refolded receptor.

2. To express, purify and characterize hNK1R C-terminus to elucidate its structural properties using spectroscopic techniques in order to deepen the knowledge of described functional roles of this domain.

The specific aims of this study are as follows:

1. To design and subclone genes of different forms of hNK1R and hNK1R C-terminus in suitable vectors containing 6X-His tag for expression in C43 strain of *E. coli*.

2. To express hNK1R-366, hNK1R-311 truncated forms and hNK1R C-terminus in *E. coli* in form of inclusion bodies and screen for an appropriate detergent for their solubilization.

3. To find suitable conditions for refolding of hNK1R-366 and hNK1R-311 and to determine their capacity for binding SP using radioligand binding assay.

4. To characterize the refolded hNK1R-366 and hNK1R-311 by different spectroscopic techniques.

5. To determine the secondary structure of hNK1R C-terminus using Circular Dichroism and FTIR spectroscopic techniques.

6. To study by means of UV and fluorescence spectroscopy techniques the intrinsic conformational changes in the tertiary structure of hNK1R C-terminus in order to deepen the knowledge of the structure and functional roles of this domain.

7. To propose, according to the spectroscopic experimental results, a model for tertiary structure of hNK1 C-terminus and to indicate its most important characteristics.

MATERIALS AND METHODS

3

3.1. Chemicals, solutions, kits and laboratory equipment

3.1.1. Chemicals and kits

Tris/Glycine/SDS 10X Buffer	Biorad
T4 DNA ligase,	Roche (Germany)
DNAseI	Roche (Germany)
Nonidet-P40	Roche (Germany)
Restriction enzymes	New England Biolabs
IPTG	Apolloscientific (UK)
Magnesium chloride	Callbiochem (Merck, Germany)
Ampicillin	Melford (UK)
AEBSF	Melford (UK)
Sodium chloride	Melford (UK)
Sodium phosphate monobasic	Melford (UK)
Sodium phosphate dibasic	Melford (UK)
Glycerol	Melford (UK)
Imidazole	Melford (UK)
Guanidinium hydrochloride	Melford (UK)
Sodium dodecyl sulfate (SDS)	Melford (UK)
Tris base,	Melford (UK)
n-Dodecyl *β*-D-maltoside (DDM)	Melford (UK)
Laurylamido-*N*,*N*'-Dimethylpropyl Amine Oxide	Anatrace (OH, USA)
N-Lauroylsarcosine (Sarcosyl)	Fluka (Sigma-Aldrich, MO, USA)
Sodium deoxycholate	Sigma Aldrich (MO, USA)
Sucrose	Sigma Aldrich (MO, USA)
Lysozyme	Sigma Aldrich (MO, USA)
Triton-X100	Sigma Aldrich (MO, USA)
His-select nickel affinity gel	Sigma Aldrich (MO, USA)
Rabbit polyclonal to 6X his tag (HRP) antibody	Abcam
Oligonucleotide primers for PCR	Invitrogen
KOD Hot Start DNA polymerase kit	Novagen
Agarose	Melford (UK)
QIAquick PCR Purification Kit	Qiagen
QIAGEN Plasmid Purification Kit	Qiagen
QIAprep Miniprep kit	Qiagen
Dubelcco's Modified Eagle Medium (DMEM)	Gibco
Calf fetal bovine serum (FBS)	Gibco
Yeast extract	Pronadisa
Tryptone	Pronadisa
Penicillin/Glutamine/Streptomycin solution	Gibco

3.1.2. Laboratory equipment

Sonic dismembrator Dyntech
Orbital shaker Heidolph model Polymax 1040
Rotating shaker Atom 85
Autoclave Stericlav model 75
Balances
Termostatic bath Tectronbio
Termic block Labline model Multiblock
Vaccum pump Vacuumbrand model M2 2C NT
Biological safety laminar flow cabinet Nuaire model 425-400E, class IIA
Cell scrapers Costar model III
Table centrifuge Sorvall model TC6
Centrifuge Kontron Centrikon H-401, rotors A8.24 and A6.14
Agarose Gel Electrophoresis System SCIE-PLAS
SDS-PAGE electrophoresis System Miniprotean Tetracell Biorad
Mini Trans-Blot System Biorad
Power supplies Power Pac 200 and 300 Biorad
Incubator CO2 Nuaire Water Jacketed
Microcentrifuge Eppendorf 5418
Micropipettes Nichipet
Inverted microscope Olympus model CK2
pH-meter Metrohm model 713
Liquid nitrogen Duar vessel Air Liquide model GT35
Ultracentrifuge Sorvall model Combi Plus, rotor T-865
Vortex Heidolph model Reax 2000
FTIR spectrophotometer Varian 7000e
JASCO J-715 spectropolarimeter JASCO
Spectrophotometer Cary 50-bio Varian
Spectrofluorometer QuantaMaster 30 Photon Technology International, USA
Micro ultracentrifuge Sorvall Discovery M150
Centrifuge Beckman Avanti J-25 rotors JLA 10.500 and JA 25.50
Tissue culture plates 75 cm2 TPP
Thermo Scientific maxQ 4000
Infors AG CU4103 Bottmingen
Packard Tri-Carb Liquid Scintillation Counter

3.2. Molecular cloning and DNA manipulation techniques

3.2.1. Plasmid vectors and genes

For cloning and expression of hNK1 receptor and hNK1R C-terminus the following expression vectors were used:

pET23a

It is a transcription vector designed for expression from bacterial translation signals carried within a cloned insert. The pET23a vector lacks the ribosome binding site and ATG start codon present on the pET translation vectors. A C-end or N-end His-tag sequence is available at user's choice. Sequencing of cloned gene can be performed using T7 promotes and T7 terminator primer. We used pET23a plasmid vector for cloning and *E. coli* expression of hNK1 receptor and hNK1R C-terminus.

pRS314

This plasmid vector was used as a source of human full-length NK1receptor gene and was kindly donated by Dr. Anne Robinson from the Department of Chemical Engineering of Delaware State University.

pCDNA3

pDNA3 is a 5.4 kb vector designed for high-level stable and transient expression in mammalian hosts under human cytomegalovirus immediate-early (CMV) promoter. The vector has a multiple cloning sites in the forward (+) and reverse (-) orientations to facilitate cloning, neomycin resistance gene for selection of stable cell lines, episomal replication in cells lines that are latently infected with SV40 or that express the SV40 large T antigen (e.g. COS-1, COS-7) and gene of resistance to ampicillin for *E. coli* colonies selection. We used pCDNA3 vector for hNK1 receptor expression in COS-1 cell line. *Nde*I and *Xho*I cleavage sites were used for gene cloning.

3.2.2. DNA fragment amplification by PCR

KOD Hot Start DNA Polymerase, a premixed complex of a KOD HiFi DNA Polymerase and two monoclonal antibodies was used for PCR amplification. The KOD Hot Start couples the high fidelity, fast extension speed and processivity. It efficiently amplifies genomic and phage/plasmid DNA targets up to 12 and 20 kbp producing blunt-ended DNA products that are suitable for cloning.

- Assemble the following reaction in a 0.5 ml PCR tube at room temperature or on ice:

  5 µl 10X PCR Buffer for KOD Hot Start DNA Polymerase
  5 µl dNTPs (final concentration 0.2 mM)
  4 µl MgSO4
  0.6 µl Primer FOR 25 µM stock (final concentration 0.3 µM)
  0.6 µl Primer REV 25 µM stock (final concentration 0.3 µM)
  1 µl DNA template
  1 µl KOD Hot Start DNA polymerase (1U/ µl)
  32.8 µl PCR grade water

  Total volume: 50 µl

- Mix gently and centrifuge briefly if necessary to bring reaction components to the bottom of the tube and place in thermal cycler.
- Activate the polymerase by heating for 2 min at 94 °C followed by:

  Denature: 15 sec at 94°C
  Anneal: 30 sec at 60°C
  Extend: 20 sec / kbp at 72°C
  Repeat for 28–30 cycles

- To analyze the reaction products, remove a 5 µl sample, add 6X Loading buffer, run a 1% agarose gel containing 0.5 µg/ml ethidium bromide and visualize the bands under UV illumination.

3.2.3. Purification of PCR products

Products of PCR amplification should be purified prior to any other cloning manipulations (ligation, enzymatic digestion, etc.) We used QIAquick PCR Purification Kit designed to purify single- or double-stranded DNA fragments from PCR and other enzymatic reactions Fragments ranging from 100 bp to 10 kb can be purified from primers, nucleotides, polymerases, and salts using QIAquick spin columns in a microcentrifuge. For proper PCR products purification QIAquick PCR Purification Kit instruction manual should be used.

3.2.4. Enzymatic digestion of DNA with endonucleases

Restriction endonuclease digestion of DNA is an essential part of cloning procedure and yields DNA fragments of a convenient size for downstream manipulations. Restriction endonucleases are bacterial enzymes that bind and cleave DNA at specific target sequences. Type II restriction enzymes are the most widely used in molecular biology applications. They bind DNA at a specific recognition site, consisting of a short palindromic sequence, and cleave within this site. If a DNA fragment is to be cut with more than one enzyme, both enzymes can be added to the reaction simultaneously provided that they are both active in the same buffer and at the same temperature.

- Pipet reaction components into a tube and mix well by pipetting.

  5 μl DNA
  2 μl Restriction enzyme buffer 10X
  1 μl Endonuclease 1
  1 μl Endonuclease 2
  11 μl milliQ water

  Total volume: 20 μl

- Incubate the digest in a water bath for 1.5 h at temperature suitable for the particular enzyme.
- Heat the reaction to 65 °C for 20 min after digestion to inactivate the enzyme.

3.2.5. Mini-prep plasmid DNA purification

QIAprep Miniprep kit (Qiagen) was used to prepare a high grade purity plasmid DNA. The QIAprep Miniprep system provides a fast, simple, and cost-effective plasmid miniprep method for routine molecular biology laboratory applications. Silica membrane technology is used to eliminate the cumbersome steps associated with loose resins or slurries. Plasmid DNA is eluted in a small volume of Tris buffer and is immediately ready for use.

- Pick a single colony from a freshly streaked selective plate and inoculate a culture of 5 ml LB medium containing the appropriate selective antibiotic.
- Incubate for overnight at 37 °C with vigorous shaking. Growth for more than 16 h is not recommended since cells begin to lyse and plasmid yields may be reduced.
- Harvest the bacterial cells by centrifugation at 6,800 x *g* in a conventional, table-top microcentrifuge for 3 min at room temperature (15–25 °C).
- Proceed with purification using QIAprep Miniprep kit instruction manual.

3.2.6. Maxi-prep plasmid DNA purification

Large quantities of plasmid DNA needed for mammalian cells transfection were obtained using QIAGEN Plasmid Purification Kits which is based on the remarkable selectivity of patented QIAGEN Resin, allowing purification of ultrapure supercoiled plasmid DNA with high yields. QIAGEN plasmid purification protocol is based on a modified alkaline lysis procedure, followed by binding of plasmid DNA to QIAGEN Anion-Exchange Resin.

- Pick a single colony from a freshly streaked selective plate and inoculate a starter culture of 2–5 ml LB medium containing the appropriate selective antibiotic.
- Incubate for approximately 8 h at 37 °C with vigorous shaking.
- Dilute the starter culture into selective LB medium supplied with an appropriate antibiotic (50 mg/ml). For high-copy plasmids, inoculate 500 ml medium with 200 *µ*l of starter culture.
- Grow at 37 °C for 12–16 h with vigorous shaking (approx. 300 rpm).

- Harvest the bacterial cells by centrifugation at 6,000 x *g* for 15 min at 4 °C.
- Proceed with purification using QIAGEN Plasmid Purification Kit instruction manual.

3.2.7. DNA fragments ligation

T4 DNA ligase is used to ligate DNA fragments with sticky or blunts ends generated by enzymatic restriction o PCR.

- Pipet reaction components into a tube and mix well by pipetting.

2 μl T4 DNA ligase 10X Buffer
4 μl linearized vector
13 μl insert
1 μl T4 DNA ligase

Total volume: 20 μl

- Incubate at 16 ºC overnight (for sticky ends fragments) and at 25 ºC for 3-4 hours for blunt ends fragments.
- Transform DH5*α* competent cells with the ligation products.

3.2.8. Electrophoresis of DNA in agarose gel

For 1% agarose gel preparation:

- weight 50 mg of agarose and add 50 ml of 1X TAE buffer (40 mM TrisHCl 20 mM acetic acid, 1 mM EDTA, pH 7.6).
- Heat agarose suspension in microwave oven until it starts boiling.
- Cool down agarose solution, add 4 μl of ethidium bromide and pour the mixture into the gel rack.
- Insert the comb at one side of the gel and wait for 20 minutes until it becomes solid.
- Remove the comb and put the gel, together with the rack, into a tank filled with 1X TAE buffer.

- After the gel has been prepared, use a micropipette to inject 5 *μ*l of prestained DNA ladder into the first slot and fill other slots with 5 *μ*l of stained DNA.
- Close the lid of the electrophoresis chamber and set the current at 110 V for 40 min.
- When the colored "front wave" of DNA ladder and DNA samples approaches the end of the gel, the current is stopped.
- Visualize the DNA stained with ethidium bromide under ultraviolet light.

3.3. Production of the recombinant proteins in *E. coli*

3.3.1. General characteristics of *E. coli* strains

Two *E. coli* strains were used for cloning and expression of NK1R truncated forms and NK1R C-terminus.

DH5*α* strain:

Genotype: F- endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG *Φ*80dlacZ*Δ*M15 *Δ*(lacZYA-argF)U169, hsdR17(rK- mK+), *λ*–

DH5*α* strain can be transformed with high efficiency and like many cloning strains has several features that make it useful for recombinant DNA methods. The endA1 mutation inactivates an intracellular endonuclease that degrades plasmid DNA in many miniprep methods. The hsdR17 mutation eliminates the restriction endonuclease of the *Eco*RI restriction-modification system, so DNA lacking the *Eco*RI methylation will not be degraded. DNA prepared from hsdR strains that are wild type for hsdM will be methylated and can be used to transform wild type *E. coli* K-12 strains. *Δ*(lacZ)M15 is the alpha acceptor allele needed for blue-white screening with many lacZ based vectors. recA eliminates homologous recombination. This makes the strain somewhat sickly, but reduces deletion formation and plasmid multimerization.

C41(DE3) strain:

The strain C43(DE3) was derived from BL21(DE3) (F- ompT hsdSB (rB- mB-) gal dcm) through C41(DE3) strain. This strain has at least one uncharacterized mutation, which prevents the cell death associated with the expression of toxic recombinant proteins.

3.3.2. Culturing conditions and media

Luria Bertani (LB), a nutritionally rich medium, is primarily used for the growth of *E. coli*. There are several common formulations of LB. Generally ingredients to promote growth are used, including peptides provided by tryptone, casein peptones, vitamins, trace elements (nitrogen, sulfur, magnesium, etc.) and minerals provided by yeast extract. Sodium ions for transport and osmotic balance are provided by sodium chloride. Tryptone is used to supply essential amino acids to the growing bacteria, while the yeast extract is used to provide organic compounds helpful for bacterial growth.

LB medium

5 g yeast extract
10 g tryptone
10 g NaCl

Add the reagents to 900 ml distilled water in a measuring cylinder. Shake until all is dissolved. Adjust the final volume to 1 liter using distilled H_2O. Sterilize by autoclaving on wet cycle (20 min at 15 psi = 1.02 atm).

LB solid media

5 g yeast extract
10 g tryptone
10 g NaCl
15 g/l bacteriological grade Agar

Autoclave on wet cycle (20 min at 15 psi = 1.02 atm) and cool down carefully to 37 ºC. Add antibiotics for colonies selection (50 mg/l) and pour 20 ml of LB solid medium to each petri plastic plate. Store plates for 1 month at 4 ºC in inverted position.

3.3.3. Plasmids and cultures conservation

E. coli colonies with plasmid clones can be stored at 4 ºC during 30 days. If longer storing times are necessary, glycerol stocks should be used.

- Pick a single colony of the clone off of a plate and grow an overnight in LB medium supplemented with 50 mg/L ampicillin.
- Add 0.5 ml of the o/n culture to 0.5 ml of 80% sterile glycerol in the sterile screw cap microcentrifuge tube
- Vortex for several seconds.
- Freeze the glycerol stock at –70 ºC.

Plasmids can be stored in aqueous solution or TE buffer (10 mM Tris-HCl, 1mM EDTA, pH 8) at –20 ºC.

3.3.4. Competent *E. coli* cells preparation

- Streak *E. coli* cells on an LB plate without antibiotics.
- Allow cells to grow at 37 ºC o/n.
- Inoculate a single colony in 10 ml LB media and grow o/n at 37 ºC.
- Transfer 5 ml overnight culture into 500 ml LB media in 3 L flask.
- Allow cell to grow at 37 ºC, until $OD_{600} = 0.4$
- Transfer cells to 2 centrifuge bottles (250 ml), and place cells on ice for 20 minutes.
- Centrifuge cells at 4 ºC for 10 minutes at 3,000 x *g*. Cells must remain cold for the rest of the procedure: Transport tubes on ice and resuspend on ice in the cold room.
- Pour off media and resuspend cells in 30 ml of cold 0.1 M $CaCl_2$. Transfer the suspended cells into 50 ml polypropylene falcon tubes, and incubate on ice for 30 min.
- Centrifuge cells at 4 ºC for 10 minutes at 3,000 x *g*.
- Pour supernatant and resuspend cells (by pipetting) in 8 ml cold 0.1 M $CaCl_2$ containing 15% glycerol. Transfer 140 μl into (1.5 ml) ependorff tubes placed on ice. Freeze the cells in liquid nitrogen. Cells stored at -80 ºC can be used for transformation for up to 6 months.

3.3.5. *E. coli* heat shock transformation protocol

- Thaw an aliquot of competent cells on ice for 10 min.
- Add maximum 20 μl of a plasmid or ligation reaction and mix gently.
- Incubate the tubes on ice for 30 minutes.
- Heat shock the cells for 2 min at 42 °C water bath.
- Place the tubes immediately on ice for 2 minutes.
- Add 800 μl of LB medium without antibiotics to each tube and incubate for 1 hour at 37 °C.
- Spin down briefly and remove most supernatants.
- Resuspend cell pellet with the rest of LB medium in the tube by pipetting and plate out the suspension on a LB agar plate containing the appropriate antibiotic.
- Incubate the plates overnight at 37 °C

3.3.6. Expression of the recombinant protein in *E. coli*

C43(DE3) competent cells were transformed by thermal shock with pET23a vector carrying hNK1-366, hNK1R-311 or hNK1R C-terminus gene and plated onto solid LB medium supplemented with 50 mg/ml ampicillin for selection. A single colony was inoculated in 10 ml of LBA medium and grown overnight at 37 ºC. 16 hours later, the culture was transferred in to 100 ml of LBA and incubated at 37 ºC until the OD_{600} reached 1.0. The protein expression was induced by 2 mM IPTG and the culture was incubated for 4 additional hours. *E. coli* cells expressing hNK1R or hNK1R C-terminus were cooled down on ice for 10 min to stop the growth, pelleted by centrifugation at 16,000 × *g* for 10 min and resuspended in Buffer E (50 mM TrisHCl, 100 mM NaCl pH 8.0) (10 ml/g cells) supplemented with 25% (w/v) sucrose, 300 μM AEBSF and 4 mM TCEP to prevent the peptide from degradation by proteases and oxidation of cysteine residues.

3.3.7. Inclusion bodies extraction

Cell walls of *E. coli* bacteria resuspended in Buffer E were disrupted by adding 6 mg of lysozyme per 1 g of wet cells followed by 30 min incubation at 25 ºC. Three freeze/thaw cycles were performed to completely lyse the bacteria and 20 mM $MgCl_2$, 20 *µ*l/ml DNAseI were added to the solution to reduce viscosity caused by genomic DNA liberation. After the lysate viscosity decreased an equal volume of Buffer E supplemented with 1% (v/v) Nonidet P-40 and 1% (w/v) sodium deoxycholate was added to the lysed cells to solubilize the membranes. After 30 min of incubation at 25 ºC the sample was centrifuged at 16,000 × *g* to separate the IBs from soluble proteins. The pellet containing the IB was washed 2 times with 1% (v/v) Triton X-100 in Buffer E to eliminate non-solubilized membranes and remaining cell debris and then one time with Buffer E to remove the Triton X-100 detergent. The extracted IBs were resuspended in Buffer E supplemented with 4 mM TCEP, aliquoted and stored at -20 ºC to avoid protein aggregation.

3.3.8. IBs solubilization and purification of the recombinant proteins

The IBs obtained from 50 ml of *E. coli* culture were resuspended in 1 ml of the Buffer E supplemented with 4 mM TCEP. Distinct detergents and chaotropic agents as 6 M GuHCl, 1% (w/v) SDS, 1% (v/v) Sarkosyl, 1% (w/v) DDM and 1% (w/v) LAPAO were added to the samples to solubilize the receptor from the inclusion bodies. The solubilization was carried over for 2 h at room temperature and constant mild agitation.

The extracted IB from 1 liter of IPTG induced *E. coli* culture were resuspended in 30 ml of Buffer E by sonication and 6 M GuHCl or 1% Sarkosyl was added to solubilize the protein. hNK1R C-terminus samples that were used in FTIR assays were treated with *N*-ethylmaleimide (NEM) (1 mg of NEM / mg IBs) prior to purification in order to prevent disulfide bonds formation. TCEP cannot be used for this purpose due to its high absorption in amide I and amide II regions.

3 ml of His-select nickel affinity gel was washed 2 times with 2 volumes of miliQ water and then equilibrated for 5 min with 5 volumes of Buffer E supplemented with 6 M GuHCl or 1% Sarkosyl. 15 ml Bio-Rad chromatography column was packed

with the equilibrated His-select nickel affinity gel. 30 ml of solubilized IBs were filtered through a 0.22 *μ*M syringe filter unit and subsequently applied to the packed column at flow rate of 250 *μ*l/min. hNK1R or hNK1R C-terminus attached to nickel affinity gel was washed 2 times with 15 ml (2 volumes of packed gel) of 25 mM sodium phosphate, 100 mM NaCl (pH 7.4) buffer at 500 *μ*l/min flow rate to remove non-specifically bound proteins and GuHCl or Sarkosyl traces. 100-300 mM imidazole or 100 mM histidine in phosphate buffer was used to elute the protein from the column. From 10 to 15 fractioned elutions were made in order to obtain pure and concentrated material. Purified hNK1R or hNK1R C-terminus eluted from the nickel affinity gel was dialyzed against the phosphate buffer to remove imidazole or histidine used for elution.

3.4. Recombinant protein expression in mammalian cells

3.4.1. General characteristics and description of COS-1 cell line

The transformed COS-1 cell line was derived from the CV-1 fibroblast line, which was initiated in March, 1964 by F. C. Jensen and colleagues from the normal kidney of an adult African green monkey. COS-1 cells are propagated as a monolayer in the alpha modification of Eagle's minimal essential medium (a-MEM) supplemented with 10% fetal bovine serum or in Dulbecco's modification of Eagle's medium (figure 3.1). They grow rapidly, with an approximate population doubling time of 18 h. Typical expression experiments using COS-1 cell line involve transfection and production of recombinant proteins. The transfection can be carried out by a number of different methods, including PEI, calcium phosphate, DEAE, and lipid transfection.

COS-1 line contains T antigen, retains complete permissiveness for lytic growth of SV40, supports the replication of ts A209 virus at 40C, and supports the replication of pure populations of SV40 mutants with deletions in the early region. Since COS-1 cells are possible hosts for propagation of pure populations of recombinant SV40 virus they should be handled under laboratory containment level 2.

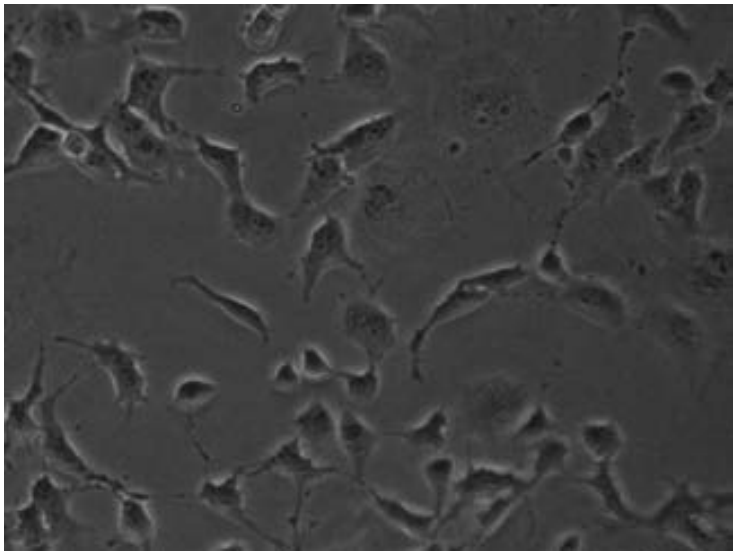


Figure 3.1 A culture of COS-1 at 50% confluency under light microscope made with 20X objective.

3.4.2. Growth and maintenance of COS-1 cell line

Complete DEMEM (Dubelcco's Modified Eagle Medium) is usually used to culture COS-1 cells. Optimal growth conditions: 37 ºC, 5% CO_2. The complete DMEM medium is usually supplemented with 10ml/L 100X Penicillin/Glutamine/Streptomycin to prevent possible bacterial contamination. Amphotericin-B 100X is added to the medium (10 ml/L) to avoid the fungal growth. The FBS has to be inactivated at 56 ºC for 30 min before adding to the medium.

Complete DEMEM (Dubelcco's Modified Eagle Medium)

Per 1 L of medium:
13.38 g DMEM powder
0.74 g $NaHCO_3$

Adjust pH to 6.8 and sterilize using Steritop sterile vacuum bottle-top 0.22 μm filtering units. Add 10% FBS; 10 ml/L Penicillin/Glutamine/Streptomycin and 10 ml/L Amphotericin-B

PBS, pH 7.2-7.3

137 mM NaCl
2.7 mM KCl
1.5 mM KH_2PO_4
8 mM $Na_2HPO4{\cdot}2H_2O$

Sterilize by autoclaving and store at 4 ºC for approximately 1 month.

3.4.2.1. Initiating the cell culture

- COS-1 cells are thawed rapidly by briefly immersing the vial in a 37 °C water bath with constant agitation. Upon thawing, immediately wipe the outside of the vial with 70% ethanol, and then transfer the contents of the vial to a 75 cm^2 culture flask.
- Add an additional 20-25 ml of the complete DMEM medium to the flask. Gently rock or swirl the flask/plate to distribute cells evenly over the growth surface. Place the culture in a 37 °C, 5% CO_2 humidified incubator.
- The next day, the cells have to be examined under a microscope. Healthy cells display a flat morphology and adhere well to the plate. Aspirate the medium and replace with fresh, pre warmed growth medium.
- Expand the culture as needed. Cell cultures should be split every 2–4 days, when they reach 70–80% confluency. As a general rule, COS-1 cells should not be allowed to become confluent nor should they be seeded too sparsely.

3.4.2.2. Preparing frozen cultures of COS-1 cells

Once COS-1 cells are established in culture, it is strongly recommended to prepare frozen stock from an early passage to ensure a renewable source of cells.

Freezing Medium:

70% DMEM
20% FBS
10% DMSO

- Detach cells from the desired number of flasks.
- Pool cell suspensions together, count cells, and calculate total viable cell number.
- Centrifuge cells at 125 x *g* for 10 min. Aspirate the supernatant.
- Resuspend the pellet at a density of at least 1–2 x 10^6 cells/ml in freezing medium.
- Dispense 1 ml aliquots into sterile cryovials.
- Freeze at -20 °C for 1 hour and then transfer the cells to –70 °C.
- Overnight remove vials from -70 ºC and place the vials in liquid nitrogen for storage.

3.4.2.3. Splitting COS-1 cells

- Aspirate the medium from the flask.
- Wash the cells carefully with PBS to remove residual medium.
- Add 10 ml of 1X trypsin in PBS and detach the cells pipetting up and down 3-4 times. Ensure that the monolayer is completely detached from the flask.
- Add 20 ml of DMEM to inactivate trypsin and centrifuge cells for 2-3 minutes at 1,500 x *g* and discard the supernatant
- Resuspend cells in 10 ml complete DMEM medium
- Take a fraction of the cell solution and inoculate the new flask. Typically, when splitting confluent COS-1 cells in a 1:10 ratio, confluency is reached again after 2-3 days.

3.4.3. Transfection of COS-1 cells using PEI method

Polyethylenimine (PEI) is a fast, effective and inexpensive reagent for large scale transient transfection. This method has been shown to express recombinant proteins in various media using COS-1 cultures. PEI condenses plasmid DNA into positively charges particles which interact with the anionic cell surface. After entering the cell through endocytosis, the high charge density of the polymer causes lysosomal rupturing and releases DNA into the cytosol.

- Dissolve PEI powder to a concentration of 2 mg/ml in water which has been heated to 80 ºC.
- Allow solution to cool to room temp and adjust pH to 7.0 with 5 M HCl.
- Filter sterilize and freeze aliquots at –80 ºC.
- Seed cells a day prior to transfection in 75 cm^2 culture flasks at a density of 2 x 10^6 cells per flask.
- Prepare the transfection solution as follows (per flask):

add 20 *µ*g of desired plasmid vector to 1300 *µ*l of basal DMEM medium (containing no FBS, antibiotics, etc.) and mix well. Add 200 *µ*l PEI (1 mg/ml) and mix by pipetting up-down. Incubate for 20 min at RT.

- Apply mixture to cells.
- Incubate the cells for 6 hour at 37 ºC and 5% CO_2.
- Wash cells with PBS and add 16 ml of complete DMEM medium.
- Incubate for 48 h and check for the protein expression.

3.5. Electrophoretic characterization of recombinant proteins

3.5.1. Electrophoresis in polyacrylamide gels (SDS-PAGE)

Sample preparation

Three microliters of 5X sampling Laemmli buffer (0.3 M TrisHCl, pH 6.8, 0.01% bromophenol blue, 10% SDS, 25% *β*-mercaptoethanol, 50% glycerol) were added to 15 *μ*l of each sample. The samples were not boiled prior to loading to the gel due to a special ability of the membrane proteins to form aggregates in SDS at high temperatures.

SDS-PAGE gel preparation

- Place the casting frame upright with the pressure cams in the open position and facing forward on a flat surface.
- Select a spacer plate of the desired gel thickness and place a short plate on the top of it.
- Orient the spacer plate so that the labeling is "up". Slide the two glass plates into the casting frame. Ensure that both plates are flush on a level surface and that the labels on the spacer plate are oriented correctly. Leaking may occur if the plates are misaligned or oriented incorrectly.
- When the glass plates are in place, engage the pressure cams to secure the glass cassette sandwich in the casting frame. Check that both plates are flush at the bottom.
- Place the casting frame into the casting stand by positioning the casting frame onto the casting gasket while engaging the spring-loaded lever of the casting stand onto the spacer plate.

- Place a comb completely into the assembled gel cassette. Mark the glass plate 1 cm below the comb teeth. This is the level to which the resolving gel is poured. Remove the comb.
- Prepare the resolving gel monomer solution by combining all reagents as described in the table below. Add APS and TEMED to the solution and pour to the mark using a glass or disposable plastic pipet. Pour the solution smoothly to prevent it from mixing with air.

Table 3.1 Composition of 12%, 18% and stacking SDS-PAGE gels for Biorad Miniprotean Tetracell SDS-PAGE electrophoresis system

Component	12 %	18 %	Stacking gel
milliQ water	4.4 ml	3 ml	1.5 ml
Acrylamide/Bis-Acrylamide Solutions 40% (w/v)	3 ml	4.4 ml	250 µl
TrisHCl 1.5 M (pH 8.8)	2.5 ml	2.5 ml	–
TrisHCl 0.5 M (pH 6.8)	–	–	250 µl
SDS 10% (w/v)	100 µl	100 µl	20 µl
APS 10%	100 µl	100 µl	20 µl
TEMED	4 µl	4 µl	2 µl

- Immediately overlay the monomer solution with water or 2-propanol.
- Allow the gel to polymerize for 45 minutes to 1 hour. Rinse the gel surface completely with distilled water. Do not leave the alcohol overlay on the gel for more than 1 hour because it will dehydrate the top of the gel.
- Prepare the stacking gel monomer solution. Combine all reagents and pour the solution between the glass plates. Continue to pour until the top of the short late is reached.
- Insert the desired comb between the spacers starting at the top of the spacer plate, making sure that the tabs at the ends of each comb are guided between the spacers. Seat the comb in the gel cassette by aligning the comb ridge with the top of the short plate.
- Allow the stacking gel to polymerize for 30–45 minutes.

- Gently remove the comb and rinse the wells thoroughly with distilled water or running buffer.
- Fill the assembly (upper chamber) with buffer to just under the edge of the outer gel plate.
- Load the samples into the wells with a Hamilton syringe or a pipet using gel loading tips.
- Place the tank on a flat surface, with the front of the tank facing you, and place the electrode assembly in the back position of the cell, making sure that the red (+) electrode jack matches the red marking on the top right inside edge of the tank.
- Fill the lower chamber with buffer to the indicated level.
- Place the Lid on the Mini-PROTEAN Tetra Tank. Make sure to align the color coded plugs and jacks.
- Insert the electrical leads into a suitable power supply with the proper polarity and adjust the voltage.

12% polyacrylamide gels were cast for hNK1R and 18% of polyacrylamide was used for hNK1 C-terminus for SDS–PAGE and Western blots analysis. The gels were run at 110 V for 1.5 h in SDS running buffer (25 mM TrisHCl, 192 mM glycine, 1% SDS, pH 8.3).

Gels staining and imaging

The gels were stained by Coomassie staining solution (10% (v/v) acetic acid, 50% (v/v) methanol, 0.25 mg/ml Coomassie blue) for 4 hours and then distained overnight with 50% (v/v) methanol, 10% (v/v) acetic acid solution. Images of the stained gels were digitized at a resolution of 300 dpi using an HP flatbed scanner.

3.5.2. Immunodetection of proteins (Western blot)

- Pre-wet materials in transfer buffer (25 mM TrisHCl, 192 mM glycine, 20% (v/v) methanol, pH 8.2). Stack in the following order (figure 3.2).
- Place the cassette in the transfer apparatus with black side facing black.
- Set-up the power supply unit at 300 mA and run transfer for 1 hour.

Figure 3.2 Western Blot transfer stack.

- Block the membrane for 30 min in 20-30 ml 1x TTBST buffer (20 mM TrisHCl, 150 mM NaCl, 0.1% Tween, pH 7.6) + 3% BSA in a small Tupperware dish on a shaker.
- Incubate with primary antibody diluted in 5 ml 1X TTBST buffer 0.1% BSA for at least 1 hour at RT or o/n at 4 ºC in a sealed bag. The primary antibody mix can be re-used (store at 4 ºC.)
- Wash the membrane 3 times for 5-10 min in ~50 ml 1X TTBST buffer.
- Incubate with secondary antibody diluted in 5ml 1X TTBST buffer + 0.1% BSA for 30 min to 1 hour at RT in a sealed bag.
- Wash 3 times, 10 min each in ~50 ml 1X TTBST buffer at RT in a small Tupperware on a shaker.
- Rinse with milliQ water.
- Detect protein with ECL kit (2 ml/membrane). In a separate tube, mix black and white ECL solutions in a 1:1 ratio. Then aliquot solution onto membranes and wait for 1 minute. Drain the ECL, wrap in plastic and expose to film.

The proteins were detected using a 6X HIS tag primary antibody conjugated with HRP (abcam) at a 1:25,000 dilution in TTBST buffer. Immunoreactive bands were detected using SuperSignal West Pico Chemiluminescent Substrate Kit (Thermo Fisher Scientific, IL, USA) and imaged with a Curix 60 CURIX 60 is a tabletop processor (AGFA).

3.6. Secondary and tertiary structure prediction

3.6.1. The use of Phyre server for secondary structure prediction

Phyre is a WEB-based server for protein secondary and tertiary structure prediction. It uses a library of known protein structures taken from the Structural Classification of Proteins (SCOP) database and augmented with newer depositions in the Protein Data Bank (PDB).The sequence of each of these structures is scanned against a non-redundant sequence database and a profile constructed and deposited in the 'fold library'. The known and predicted secondary structure of these proteins is also stored in the fold library.

Primary amino acid protein sequence should be given to Phyre server as an input. 1-letter code is used to designate each amino acid residue. The user must specify its valid e-mail address where the prediction results will be sent.

For secondary structure prediction Phyre server uses three different programs: Psi-Pred (McGuffin et al., 2000), SSPro (Pollastri et al., 2002) and JNet15 (Cole et al., 2008). The use of multiple protocols for prediction leads to more accurate interpretation of the results. The output of each program is in the form of a three-state prediction: alpha helix (H), beta strand (E) and coil (C). Each of these three programs provides a confidence value at each position of the query for each of the three secondary structure states. These confidence values are averaged and a final, consensus prediction is calculated and displayed beneath the individual predictions (Kelley and Sternberg, 2009).

3.6.2. Tertiary structure prediction by I-TASSER

I-TASSER is a hierarchical protein structure modeling approach based on the secondary-structure enhanced Profile-Profile threading Alignment (PPA) and the iterative implementation of the Threading ASSEmbly Refinement (TASSER) program (figure 3.3). The target sequence is threaded through a representative PDB structure library to search for the possible folds. The continuous fragments are then excised from the threading aligned regions which are used to reassemble full-length models while the threading unaligned regions are built by *ab initio* modeling. The conformational space is searched by replica-exchange Monte Carlo simulations. The final full-length models are generated by docking the model of

domains together. The domain docking is performed by a quick Metropolis Monte Carlo simulation where the energy is defined as the RMSD of domain models to the full-chain model plus the reciprocal of the number of steric clashes between domains (Zhang, 2008).

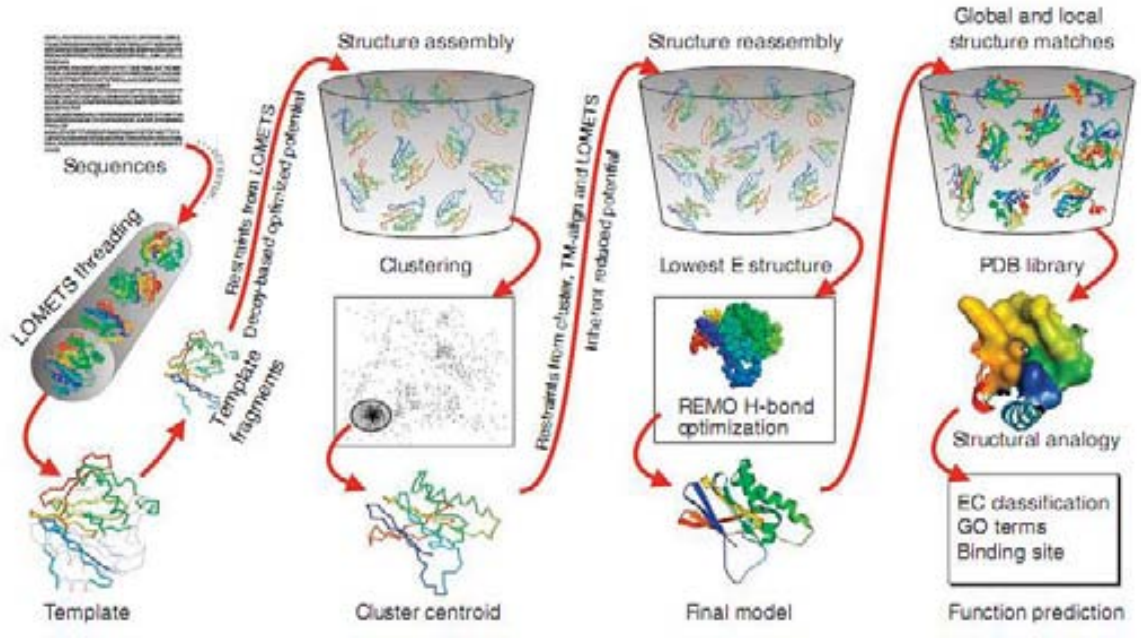


Figure 3.3 A schematic representation of the I-TASSER protocol for protein structure and function predictions. The protein chains are colored from blue at the N-terminus to red at the C-terminus. (from Roy et al., 2010)

The quality of models generated by I-TASSER can be estimated using C-score and TM-score. C-score is defined based on the quality of the threading alignments and the convergence of the I-TASSER's structural assembly refinement simulations. When a C-score cutoff of -1.5 is used to select models of correct topology, both the false-positive and the false-negative rates are below 0.1, which means that more than 90% of the quality predictions are correct. TM-score is defined to assess the topological similarity of two protein structures TM score stays in [0, 1] with higher values indicating better models. Statistically, a TM-score ≤ 0.17 corresponds to a similarity between two randomly selected structures from the PDB library; a TM-score > 0.5 corresponds approximately to two structures of the similar topology. One advantage of the TM-score is that the meaning of the TM-score cutoffs is independent of the size of proteins (Zhang, 2008; Roy et al., 2010).

The amino acid sequence of hNK1R C-terminus was provided in the FASTA format. The modeling procedure typically takes 5–10 hours for a sequence around 200 residues. After the modeling is finished, up to 5 predicted models are generated by the server, C-score of the models, and the predicted RMSD and TM-score of the first model were also provided.

The PDB files and the visual files are kept on our server for 365 days and made publicly downloadable at, so that other users can quickly retrieve the modeling results without resubmitting the jobs when they want to model the same or similar proteins. The received models were visualized and analyzed in RasTop 2.2 (Glaxo Wellcome Research & Development, Stevenage, Hertfordshire, UK) molecular visualization software adapted from the program RasMol.

3.7. Spectroscopic characterization of recombinant proteins

3.7.1. UV-vis spectroscopy

The absorption spectra of purified recombinant proteins were measured with a Cary 50-bio spectrophotometer (Varian) at scan rate 600 nm / min, average time 0.1 sec and data interval of 1 sec. All measurements were carried out 120 μl quartz Ultra-Micro cell (Hellma, Germany) and the recorded spectra were corrected for the buffer absorption by baseline subtraction.

3.7.2. Circular dichroism

All samples for CD measurements were prepared in 25 mM sodium phosphate, 100 mM NaCl (pH 7.4) buffer supplied with 4 mM TCEP to prevent the peptide oxidation and improper disulfide bonds formation. Solid GuHCl was added at different concentrations (ranging from 0 to 6 M) to denature the peptide. CD spectra of hNK1 C-terminus were recorded at room temperature using a JASCO J-715 spectropolarimeter (JASCO, Japan) in far ultraviolet region from 190 to 260 nm using the following parameters: scan speed: 10 nm/min; bandwidth 0.5 nm; cuvette path length: 0.1 cm. All spectra were corrected for background by subtraction of the buffer spectra. Mean residue molar ellipticity (θ_{MRE}) was calculated as follows:

$$\theta_{MRE} = \theta \frac{MRW}{10 \cdot l \cdot c},$$

where MRW is mean residue weight ($MW_{protein}$ / $N_{residues}$ - 1), *l* is a path length in cm and *c* is the protein concentration in mg/ml. The percentage of secondary structure elements was done according to Woody and Sreerama (Sreerama and Woody, 1993) using SELCON3 method.

3.7.3. Fluorescence spectroscopy

Fluorescence emission spectra were measured for hNK1R C-terminus in a QuantaMaster 30 Spectrofluorometer (Photon Technology International, USA). The samples were excited at 280 nm and the emission range from 290 to 400 nm was recorded. The experiment was carried out in 25 mM sodium phosphate, 100 mM NaCl (pH 7.4) buffer at room temperature. All the obtained spectra were corrected for the buffer emission.

3.7.4. ATR Fourier transform infrared spectroscopy (FTIR)

A sample of 200 *μ*l protein solution (0.2 mg/ml) in 25 mM phosphate buffer (pH 7.4) was spread homogeneously on a ZnSe crystal (Dave et al., 2001). Constant flow on nitrogen gas (220 ml/min) was applied to the crystal until the dehydrated film was formed and the ATR device was covered with a stainless steel plate containing a gas input and output to avoid the sample rehydration. ATR-FTIR scans were performed at room temperature using Varian 7000e FTIR spectrophotometer (Varian) equipped with a cooled liquid nitrogen mercury-cadmium-telluride detector. Two thousand scans were collected from 0 to 7900 cm^{-1} at a nominal resolution of 2 cm^{-1}. The recorded spectrum was corrected against the ZnSe cell background. The recorded spectra were manipulated using Varian Resolutions software. The spectra deconvolution was done and the secondary structure was estimated by band assignment of deconvolved amide I region (1700 to 1600 cm^{-1}) using the following frequency ranges in H_2O: 1652-1662 cm^{-1} for *α*-helix; 1620-1645 cm^{-1} for *β*-sheets; 1662-1690 cm^{-1} for *β*-turns and 1645-1652 cm^{-1} for random coil (Byler and Susi, 2004).

3.8. Saturation radioligand binding assay

Saturation radioligand binding experiments measure specific radioligand binding at equilibrium at various concentrations of the radioligand. Analysis of these data can be used to determine receptor number and affinity. The analyses depend on the assumption that you have allowed the incubation to proceed to equilibrium. This can take anywhere from a few minutes to many hours, depending on the ligand, receptor, temperature, and other experimental conditions. The lowest concentration of radioligand will take the longest to equilibrate. When testing equilibration time, therefore, use a low concentration of radioligand (from 10-20% of the K_d).

Binding data generated from saturation binding assay is analyzed with the equation for “total and non-specific binding” in Graph Pad Prism 4.0 software.

$$B_{L^*} = R_T \times \frac{K_L L^*}{1+K_L L^*} + NSL^*,$$

where R_T is the total receptor concentration, K_L is the radioligand equilibrium association constant, L* is the free radioligand concentration, B_L* is the bound radioligand concentration and NS is the fraction of nonspecific binding.

3.8.1. Saturation radioligand binding assay in intact COS-1 cells

- Prepare a sufficient amount of [^{3}H]-Substance P at 40 nM concentration.
- Transfect COS-1 cells with 20 μg pCDNA3-hNK1R-full using PEI at 1:6 ratio. Incubate at 37 ºC and 5% CO_2 for 48 h.
- Seed 2 x 10^6 cells in each well of 24-well culture plate. Wait for 3 h until cells get attached to the surface.
- Wash each well twice with 1 ml of cold PBS
- For total binding incubate 2 x 10^5 transfected cells at various concentration of [^{3}H]-Substance P (ranging from 0.1 to 6 nm) in Buffer RLB (50 mM TrisHCl, 1mM $MgCl_2$, 1mM $MnCl_2$, pH 7.4) for 1 h 30 min at 4 ºC. To determine non-specific binding add 1000 fold of non-labeled substance P to the mixture. Each sample should be duplicated to minimize errors that can arise from pipetting.

- After incubation wash each well with 1 ml Buffer RLB.
- Add 200 μl of 1 M NaOH to each well and transfer the lysed cells into counting vials.
- Add 200 μl of 1 M HCl to each well and carefully recover the rest of the lysed cells into counting vials.
- Add 3 ml of scintillation liquid to each counting vial and count using Tri-Carb Liquid Scintillation Counter
- Plot and analyze binding data generated from saturation binding assay in Graph Pad Prism 4.0 software. Calculate K_d and B_{max} values for the receptor. Express receptor concentration in fmol / 2 x 10^5 cells.

3.8.2. Saturation radioligand binding assay for refolded hNK1R

- Refold hNK1R on His-select nickel affinity column. Dilute 200 μl of the resin in 1 ml of the refolding buffer.
- Take a 200 μl aliqot of the diluted resin and centrifuge at 5,000 × *g* for 3 min. Discard the supernatant and add 200 μl of 200 mM imidazole to the pellet. Incubate for 15 min.
- Centrifuge the sample at 5,000 × *g* for 3 min and record a UV spectrum of the supernatant to measure Abs at 280 nm. Calculate the receptor concentration using at molar extinction coefficient at 280 nm. Determine the amount of the receptor in ng/μl of resin.
- For total binding incubate the amount of resin equivalent to 100 ng of bound receptor at various concentration of [^{3}H]-Substance P (ranging from 0.1 to 6 nM) in the Refolding buffer for 1 h 30 min at 4 ºC. To determine non-specific binding add 1000 fold of non-labeled substance P to the mixture. Each sample should be duplicated to minimize errors that can arise from pipetting.
- Filter the samples through 2 ml plastic chromatography columns and wash each column with 1 ml of 50 mM TrisHCl pH 7.4; 100 mM NaCl buffer.
- Transfer the polyethylene filter from each column into a counting vial and add 3 ml of scintillation liquid.
- Count using Tri-Carb Liquid Scintillation Counter

- Plot and analyze binding data generated from saturation binding assay in Graph Pad Prism 4.0 software. Calculate K_d and B_{max} values for the refolded receptor. Express receptor concentration in fmol.

4

RESULTS AND DISCUSSION

4.1. Forms of Neurokinin-1 receptor and the subcloning procedure

The present work comprises the study of several forms of Neurokinin-1 receptor, expressed in *E. coli*. The hNK1R-366 truncated form of the receptor ending with Glu^{366} was already described in the works of S. Bane (University of Delaware, DL, USA) and some unsuccessful refolding intents had been made (Bane et al., 2007). Since we had at our disposal the pET23b vector with subcloned hNK1R-366 ready to be expressed in *E. coli* we decided to focus our research on refolding of this particular truncated form of NK1 receptor. Furthermore, we preferred the hNK1R-366 to the full-length form hence its long C-terminus could prevent the correct folding of the receptor during the refolding process. The hNK1R-311 truncated form, ending with Arg^{311} residue also found its use in our studies. It was reported as a naturally occurring variant of the NK1 receptor in some living organism (Tuluc et al., 2009). The full-length human NK1R of 407 amino acid expressed in the membrane of COS-1 cells was used in this study as a reference in radioligand binding and some spectroscopic assays.

Another part of this work was focused on hNK1R C-terminus (the last 96 residues of the receptor from Phe^{312} to Ser^{407}) expression, purification and characterization to elucidate its structural properties. Circular dichroism along with FTIR spectroscopy are powerful techniques that have been used to study the secondary structure of the peptide. The presence of several tyrosine residues in the primary sequence of hNK1 C-terminus (figure 4.1) makes possible the use of UV and fluorescence spectroscopy techniques that can be applied for studying the intrinsic conformational changes in the tertiary structure in order to deepen the knowledge of described functional roles of this domain.

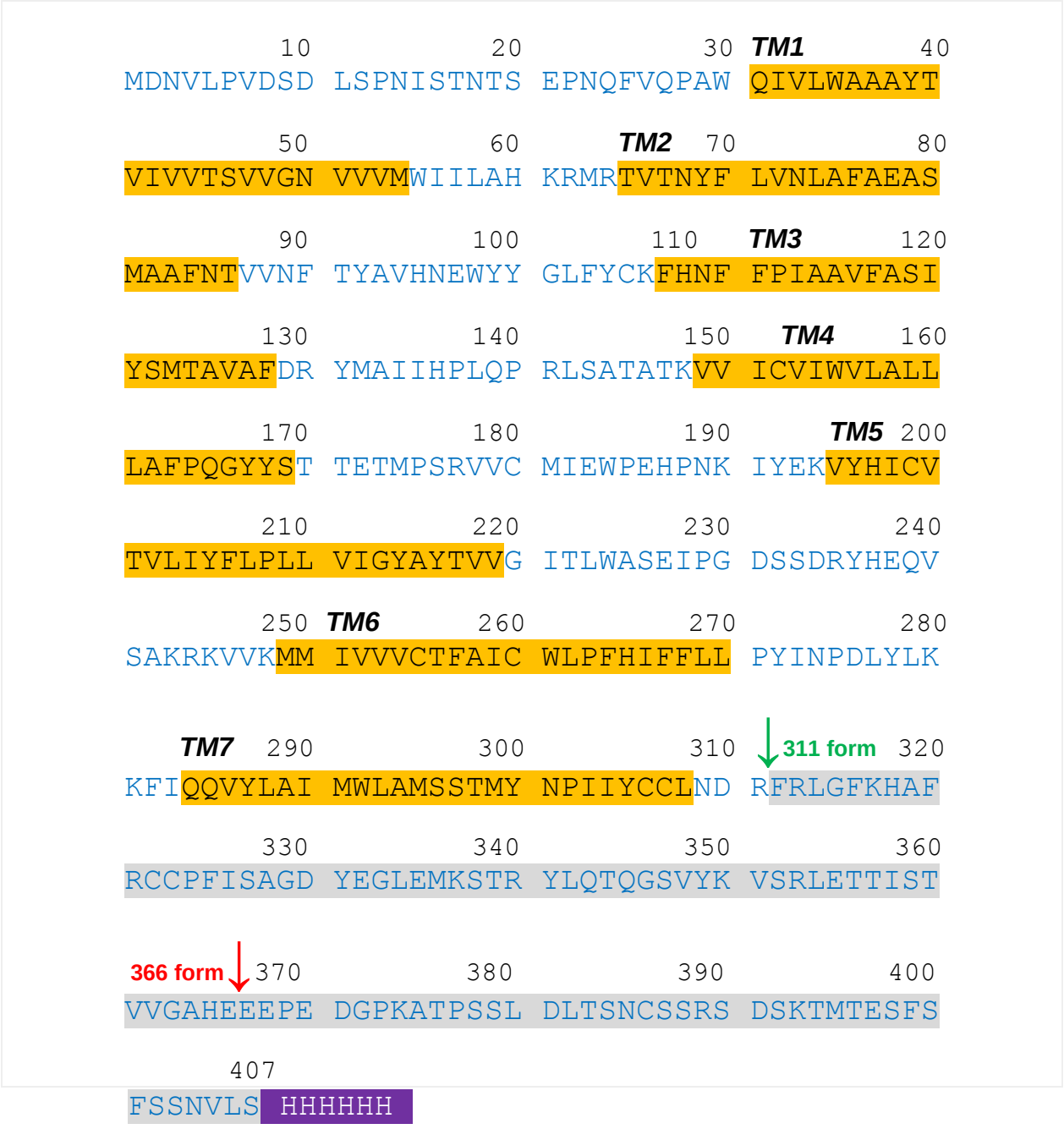


Figure 4.1 Primary amino acid sequence of human Neurokinin-1 Receptor coded by TACR1 gene. Transmembrane helices (*TM*) are shown in *orange boxes*. ↓ - hNK1R-311 truncated form; ↓- hNK1R-366 truncated form; hNK1R C-terminus is shown in *grey box*. 6X-His tag residues are shown in *purple box*.

4.1.1. Cloning hNK1R-311 truncated form into pET23a vector

pET23a is a plasmid vector designed for expression of proteins under T7 promoter in bacterial cells. We performed subcloning of truncated hNK1-311 receptor into pET23a plasmid vector for cloning and *E. coli* expression of hNK1 receptor and hNK1R C-terminus.

As a source of full-length hNK1R gene we used the pRS314 vector, kindly donated by the Department of Chemical Engineering of Delaware State University. Figure 4.2 shows the maps of both pET23a and pRS314-hNK1 vectors. The pET23a has a MSC site that allows cloning of a gene of interest between T7 promoter and T7 terminator to accomplish expression. We used *Nde*I and *Xho*I restriction sites for subcloning since *Nde*I contains an ATG start codon sequence.

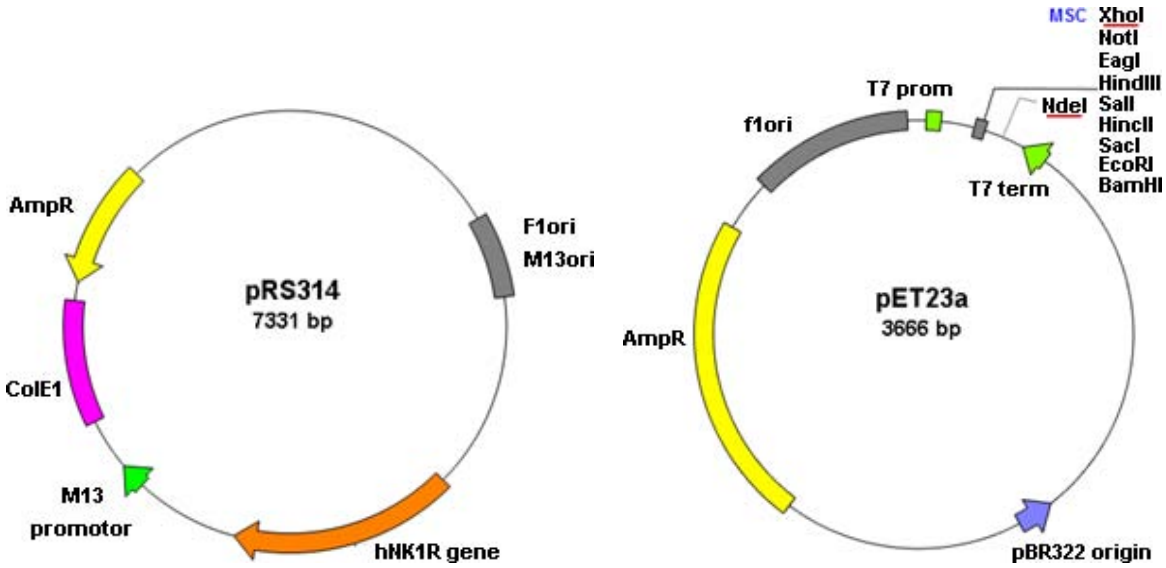


Figure 4.2 pRS314-hNK1R plasmid vector, carrying hNK1R gene and pET23a plasmid vector for protein overexpression in *E. coli*. *MSC* – multiple cloning site; cleavage sites used for subcloning are *red* and *underlined*.

The following primers containing *Nde*I and *Xho*I cleavage sites were used to amplify the hNK1R gene region from Met^{1} to Asp^{311} residues. The *hNK1-311 FOR* primer carried *Nde*I CA'TATG restriction site and the *hNK1-311 REV* primer contained a 6X-His tag placed right after Asp^{311} followed by stop codon and a *Xho*I C'TCGAG restriction site.

5’-ACCTCATATGGATAACGTCCTCCCGGT-3’ ***hNK1R-311 FOR***

5’-CAAGCTCGAGCCTGTCATTGAGGCAGCAGTAGAT-3’ ***hNK1R-311 REV***

The pET23a vector was linearized with *Nde*I/*Xho*I enzymes and the size of both fragments was verified by electrophoresis in 1% agarose gel.

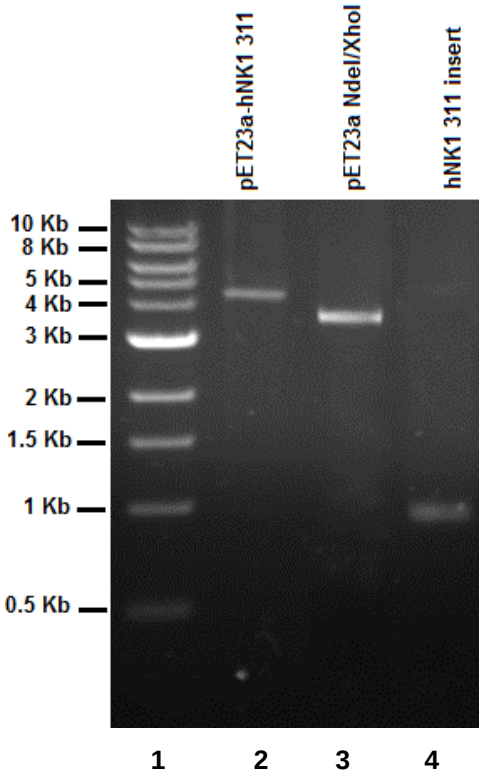


Figure 4.3 1% agarose gel stained with ethidium bromide. *Lane 1* – 1Kb DNA ladder; *lane 2* – open pET23a and hNK1R-311 insert ligation product; *lane 3* – pET23a expression vector linearized with *Nde*I/*Xho*I restriction enzymes; *lane 4* – hNK1R-311 insert.

Figure 4.3 presents a 1% agarose gel with bands corresponding to pET23a vector linearized with NdeI/XhoI (3587 pb) hNK1R-311 insert (around 950 bp). The ligation product of the open vector and the insert (*lane 2*) were purified from agarose gel and transformed into DH5*α* strain of *E. coli*. Screening for colonies gave positive results. Sequencing of mini prep of one of the colonies validated the right hNK1R-311 gene placement within the pET23a vector.

4.1.2. Cloning full-length hNK1R and truncated hNK1R-366 into pCDNA3 vector for mammalian expression

The full-length hNK1R and hNK1-366 were subcloned into pDNA3 to promote high-level transient expression in COS-1 mammalian cell line under human cytomegalovirus immediate-early (*CMV*) promoter. The pCDNA3 vector shown in figure 4.4 has a multiple cloning site (*MCS*) to facilitate cloning and an ampicillin resistance gene (*AmpR*) for selection of colonies.

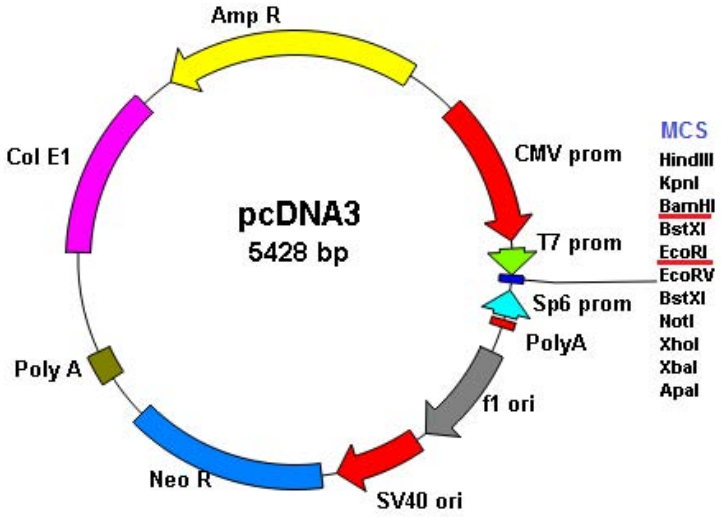


Figure 4.4 pCDNA plasmid vector for mammalian expression. *MSC* – multiple cloning site; cleavage sites that were used for subcloning are *red* and *underlined*.

The hNK1R-366 gene insert was obtained by cutting pET23b-hNK1R-366 vector with *Bam*HI/*Eco*RI enzymes. Full-length hNK1 insert was first amplified from pRS314-hNK1R template using the following primers:

5'-CTGGATCCAGTATGGATAACGTCCTCCCGG-3'	***NK1-BamHI FOR***
5'-AGCTCGAGTCAGTGGTGGTGGTGGTGGTGG GAGAGCACATTGGAGGAGAA-3'	***CMN2 REV***

pCDNA3 vector was digested with *EcoR*I/*BamH*I enzymes for hNK1-366 cloning and with *Bam*HI/*Xho*I for cloning of full-length hNK1 receptor (figure 4.5).

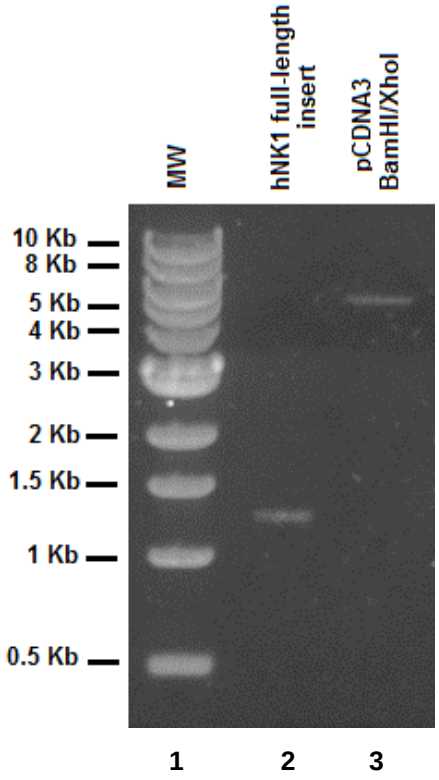


Figure 4.5 1% agarose gel stained with ethidium bromide. *Lane 1* – 1Kb DNA ladder; *lane 2* – full-length hNK1R insert; *lane 3* – pET23a expression vector linearized with *Nde*I/*Xho*I restriction enzymes.

Both inserts were ligated using T4 DNA ligase and the products of ligation were transformed into DH5*α* strain of *E. coli.* Sequencing results showed that both inserts were correctly placed into pCDNA3 vector and can be used for expression in mammalian host cells.

4.1.3. Cloning hNK1R C-terminus for *E. coli* expression

We used pET23a as a vector for expression of hNK1R C-terminus in C43 strain of *E. coli.* We amplified hNK1R gene region encoding the last 96 amino acid residues (from Phe^{312} to Ser^{407}) corresponding to hNK1R C-terminus using *NdeI-CtermhNK1 FOR* and *CtermhNK1-XhoI REV* primers that contained *Nde*I and *Xho*I restricted sites and an ATG translation initiation sequence as a part of *Nde*I cleavage site.

5'-TATCATATGTTCCGTCTGGGCTTCAAGCATGCCTTCCGG-3' ***NdeI-CtermhNK1 FOR***

5'-AGCTCGAGTCAGTGGTGGTGGTGGTGGTGGGAGAGCA CATTGGAGGAGAACTCGAGCTTG-3' ***NdeI-CtermhNK1 REV***

A 6X-His tag followed by stop codon was placed right after Ser^{407} residue. The PCR product corresponding to hNK1R C-terminus gene (300 bp) was gel purified, digested with *Nde*I and *Xho*I endonucleases and finally ligated with *Nde*I/*Xho*I open at MCS pET23a vector (3587 pb) overnight at 16 ºC using T4 DNA ligase (figure 4.6).

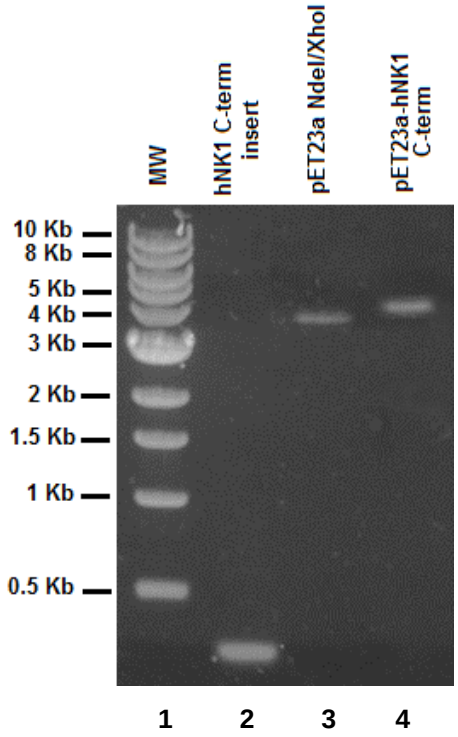


Figure 4.6 1% agarose gel stained with ethidium bromide. *Lane 1* – 1Kb DNA ladder; *lane 2* – hNK1 C-terminus insert (300 pb) digested with *Nde*I/*Xho*I endonucleases; *lane 3* – pET23a vector (3587 pb) cleaved with *Nde*I/*Xho*I; *lane 4* – pET23a-hNK1R C-terminus (3900 pb).

The ligation product (about 3900 bp) was transformed into DH5*α E. coli* strain. Two colonies were picked and inoculated in LB medium supplemented with 50 mg/ml ampicillin. The subcloned pET23a-hNK1R C-terminus vector was extracted using Qiagen Miniprep Kit (Qiagen) and subsequently verified by sequencing. The sequencing results showed correct position of hNK1R C-terminus gene within the pET23a expression vector.

4.2. Expression of different forms of Neurokinin-1 receptor in *E. coli*

Structural and biophysical studies of proteins and peptides generally require high expression levels and therefore a proper expression system should be found in every particular case. The *E. coli* expression system has been successfully used in many studies aiming to obtain large quantities of material for subsequent research. (Weiss and Grisshamer, 2002). We expressed 6X-His tagged full-length hNK1R, hNK1R-366 and hNK1R-311 truncated forms and hNK1R C-terminus in C43(DE3) *E. coli* strain. The produced proteins turned out to be insoluble in the *E. coli* cytoplasm and were expressed in form of inclusion bodies.

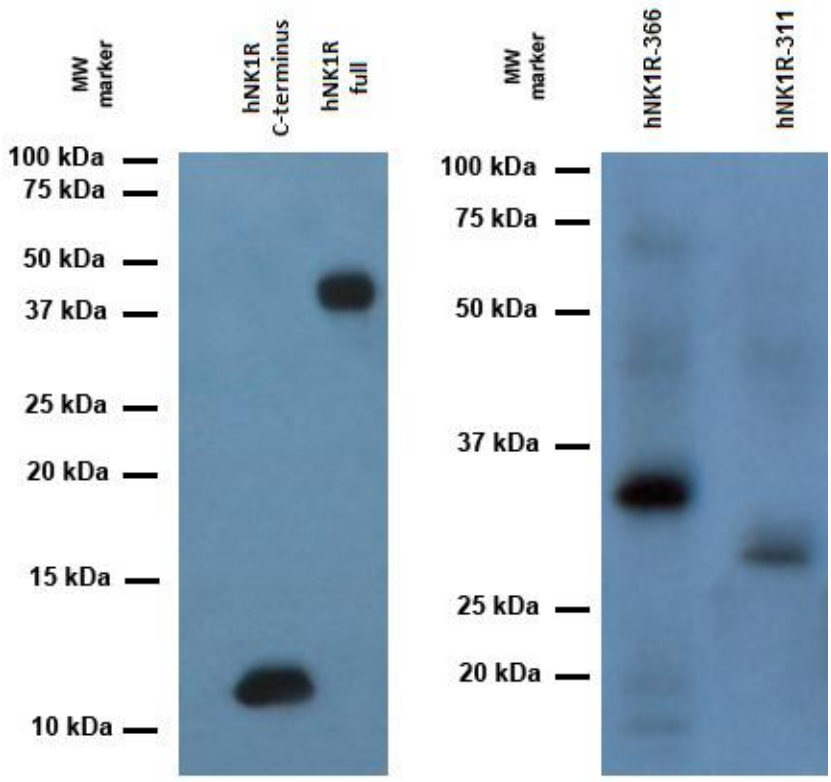


Figure 4.7 Western blot analysis of different forms of hNK1 receptor. Rabbit polyclonal antibody to 6X-His tag conjugated with HRP was used at 1:25,000 dilution to detect the proteins.

Figure 4.7 shows Western blot results for different forms of hNK1 receptor. As all expressed proteins carry 6X-His tag they can be easily detected on the WB membrane using anti 6X-His tag antibody conjugated with HRP. Full-length form of hNK1 receptor expressed in *E. coli* was only used as a size control (40 kDa band)

and not in refolding studies. Table 4.1 shows some properties of different forms of hNK1 receptor used in this work. It is worth pointing out that hNK1R-366 and hNK1R-311, when overexpressed in *E. coli* have smaller apparent molecular weight (33 kDa and 29 kDa respectively) than the calculated one. Such an altered mobility has been previously described for other membrane proteins (Turcker et al., 1996; White et al., 2004). This phenomenon could be possibly related to the fact that upon SDS denaturation 7-helix transmembrane core of membrane proteins increases its helicity acquiring greater mobility in polyacrylamide gel. Very similar behavior of hNK1R expressed in *E. coli* was described in S. Bane works. The recovered from inclusion bodies hNK1R-366 had an apparent weight lower than the theoretical one. (Bane et al., 2006)

Table 4.1 Properties of different forms of hNK1 receptor expressed in *E. coli.*

Receptor form	Number of residues	Apparent MW (kDa)	Calculated MW (kDa)	Extinction coefficient at 280 nm ($M^{-1}\cdot cm^{-1}$)
hNK1R full-length	407	42.0	47.0	83 250
hNK1R-366	366	33.0	42.8	83 130
hNK1R-311	311	29.0	35.7	79 050
hNK1R C-terminus	96	12.0	11.5	4 200

Since hNK1R C-terminus is a peptide containing only 96 amino acids, it is likely to be soluble in *E. coli* cytoplasm. However, we could detect only a small amount of hNK1R C-terminus in soluble fraction, while the insoluble fraction contained high amounts of the peptide. The hNK1R C-terminus migrates on SDS-PAGE according to its calculated molecular weight and is represented by a 12 kDa band on the WB (figure 4.7).

4.2.1. hNK1R-366 and hNK1R C-terminus IBs solubilization

Assays aiming to localize different forms of hNK1 receptor and hNK1R C-terminus expressed in *E. coli* showed that mainly they were expressed in form of inclusion bodies. The recovery of proteins usually involves the IBs extraction and their solubilization under denaturing conditions using strong detergents like SDS or

chaotropic agents at high concentrations as urea and GuHCl (Song et al., 2009). We performed a screening of different detergents and chaotropic agents on their ability to solubilize hNK1R-366 and hNK1R C-terminus inclusion bodies. SDS, Sarkosyl, DDM, LAPAO, Urea and GuHCl were tested.

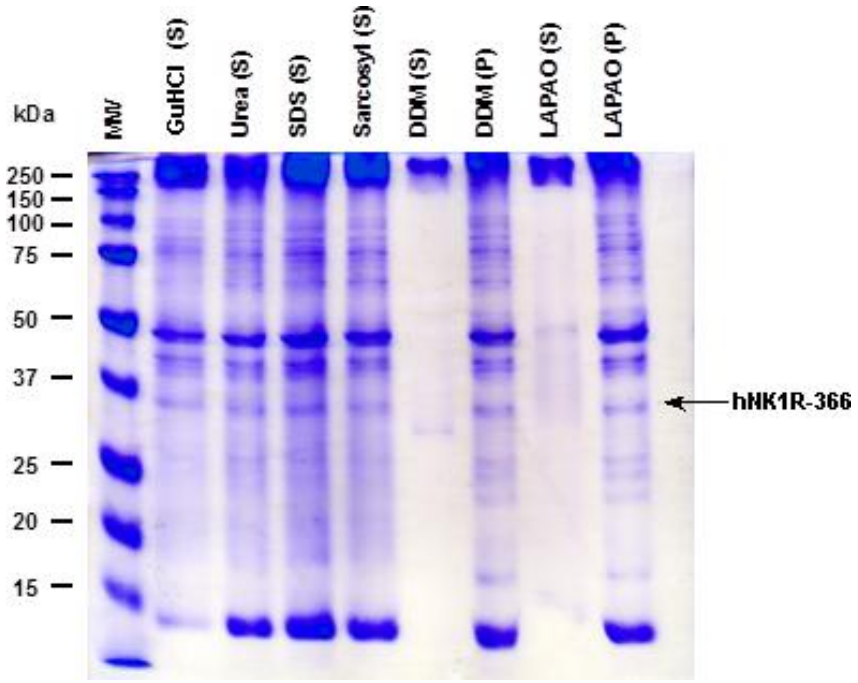


Figure 4.8 Analysis of the ability of different detergents and chaotropic agents to solubilize hNK1R-366 from inclusion bodies. SDS-PAGE 12% gel stained with Coomassie R-250 shows two fractions, the soluble (S-fraction) and a non-soluble pellet left after solubilization (P-fraction). MW is a molecular weight marker. A hNK1-366 33 kDA band is marked with an arrow on the right side of the gel.

Figure 4.8 shows SDS-PAGE of the soluble and non-soluble fractions of hNK1R-366 inclusion bodies after their solubilization. The soluble fractions for GuHCl and urea as well as SDS and Sarcosyl present a band around 33 kDA (indicated with an arrow) which corresponds to the overexpressed receptor. Mild non-ionic detergents as 1% DDM and LAPAO showed negative results as far as we were not able to detect the corresponding to NK1R-366 band on SDS-PAGE gel.

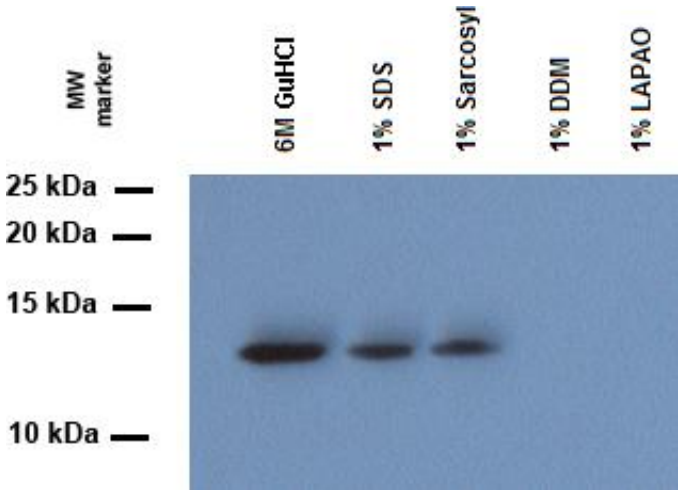


Figure 4.9 Western blot of hNK1R C-terminus expressed in C43(DE3) *E. coli* strain in form of inclusion bodies solubilized with different detergents and GuHCl. Rabbit polyclonal antibody to 6X-His tag conjugated with HRP was used at 1:25,000 dilution to detect hNK1R C-terminus. A 12 kDA band is observed when strong ionic detergents as 1% SDS, 1% sarkosyl or chaotropic 6M GuHCl are used. Mild detergents like DDM or LAPAO at 1% concentration are unable to extract hNK1R C-terminus from *E. coli* inclusion bodies.

We also tested distinct detergents on their ability to solubilize hNK1R C-terminus from *E. coli* inclusion bodies (figure 4.9). 4 mM TCEP was added to the sample to prevent the disulfide bond formation between the cysteines present in the hNK1R C-terminus. 6 M GuHCl and strong ionic detergents as 1% (w/v) SDS and sarkosyl were able to completely solubilize the peptide, whereas milder non-ionic 1% (w/v) DDM and LAPAO showed negative results. Since only strong ionic detergents or chaotropic agents are capable of extracting hNK1R C-terminus from the inclusion bodies we suggest the use of 6 M GuHCl for this purpose because it can be easily removed during purification, sample concentration and dialysis steps as compared with micelle forming detergents like SDS or Sarkosyl.

4.3. Human NK1 receptor refolding study

Membrane proteins play a fundamental role in diverse human diseases and their therapy, but suffer from a lack of structural and functional information compared to their soluble counterparts (Molina et al., 2008). *E. coli* is one of the most widely and successfully used expression systems for the production of large amounts of protein for biochemical and structural research. However, the overexpressed proteins are more likely to form inclusion bodies or the expression will be extremely low, as a result of degradation or the toxic effects of membrane protein expression in the cell. The development of strategies to overcome this barrier would significantly boost membrane protein structural research. In this regard, membrane proteins expressed as IBs are useful and could become an excellent starting point for producing membrane proteins in large quantities, provided procedures can be developed to reconstitute them in vitro (Rogl et al., 1998). Our study aims to find suitable refolding conditions for hNK1 receptor overexpressed in *E. coli* in form of inclusion bodies. We describe the use of affinity chromatography on nickel-chelating Sepharose to purify and fold complex proteins in one step. We used radioligand binding assay to determine whether the refolded receptor is active and can interact with its ligand, the SP as it occurs in natural conditions. Also, we performed the study of the refolding process itself by monitoring changes in the intrinsic fluorescence of the tryptophan residues of the protein as well as by measuring mean residue ellipticity obtained in CD experiments.

4.3.1. Purification of hNK1R-366 and hNK1R-311 from *E. coli* in denaturing conditions

An efficient purification of solubilized hNK1R C-terminus can be achieved by immobilized metal ion affinity chromatography (IMAC) via 6X-His tag on the C-end of the peptide (Zahn et al., 1997; Rogl et al., 1998). We purified hNK1-366 and hNK1-311 on His-select Nickel affinity column in denaturing conditions. The inclusion bodies solubilized in 1% Sarkosyl were injected into the column and then washed several times with Buffer E (50 mM TrisHCl, 100 mM NaCl pH 8.0) containing 1% Sarkosyl.

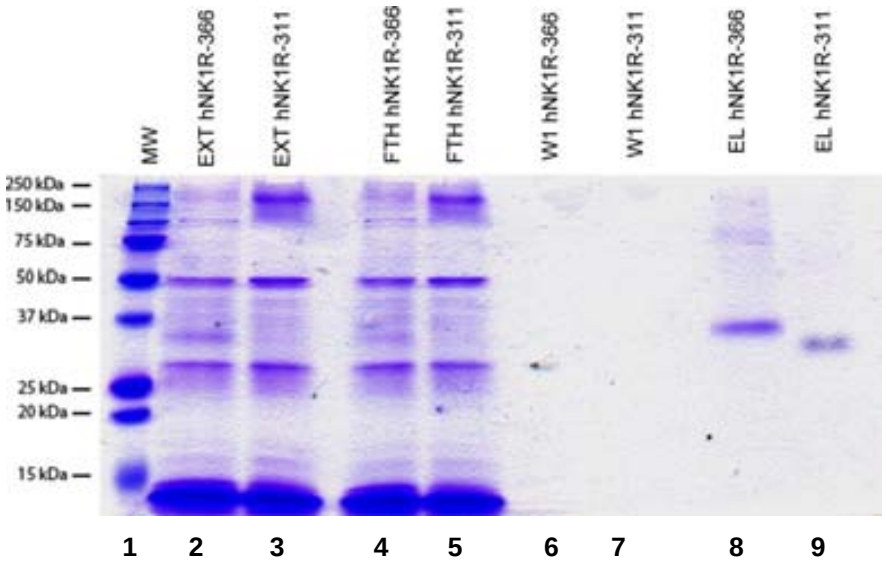


Figure 4.10 Purification results for hNK1R-366 and hNK1R-311 in 1% Sarkosyl. The samples are named as follows: *EXT* – whole inclusion bodies extract; *FTH* – flow-through of the column; *W1*, *W2* and *W3* – column washes; *EL* – protein eluted with 300 mM imidazole. 12% SDS-PAGE gel stained was stained with Coomassie solution.

The figure 4.10 shows hNK1R-366 and hNK1R-311 purified on His-select nickel affinity column in 1% Sarkosyl. As it can be noticed from SDS-PAGE gel (*lanes 2 and 3*) hNK1R-366 expression rate in *E. coli* is higher than that of hNK1-311 form of the receptor. Eluted with 300 mM receptors are shown in *lanes 8* and *9*. The Purified in 1% Sarkosyl hNK1R-366 has an apparent molecular weight of 33 kDa and hNK1R-311 29 kDa respectively.

4.3.2. Purification and on-column refolding of the hNK1R-366 and hNK1R-311 extracted from *E. coli* inclusion bodies

Refolding of membrane proteins overexpressed in *E. coli* is not a simple task. Only a few membrane proteins have been successfully refolded from inclusion bodies. On-column refolding appears to be a powerful approach for solving difficult refolding problems and can be more efficient than refolding in solution. The approach of refolding membrane protein IBs while immobilized on chelating resins has been shown to be useful for both *α*-helical and *β*-barrel membrane proteins (Rogl et al., 1998).

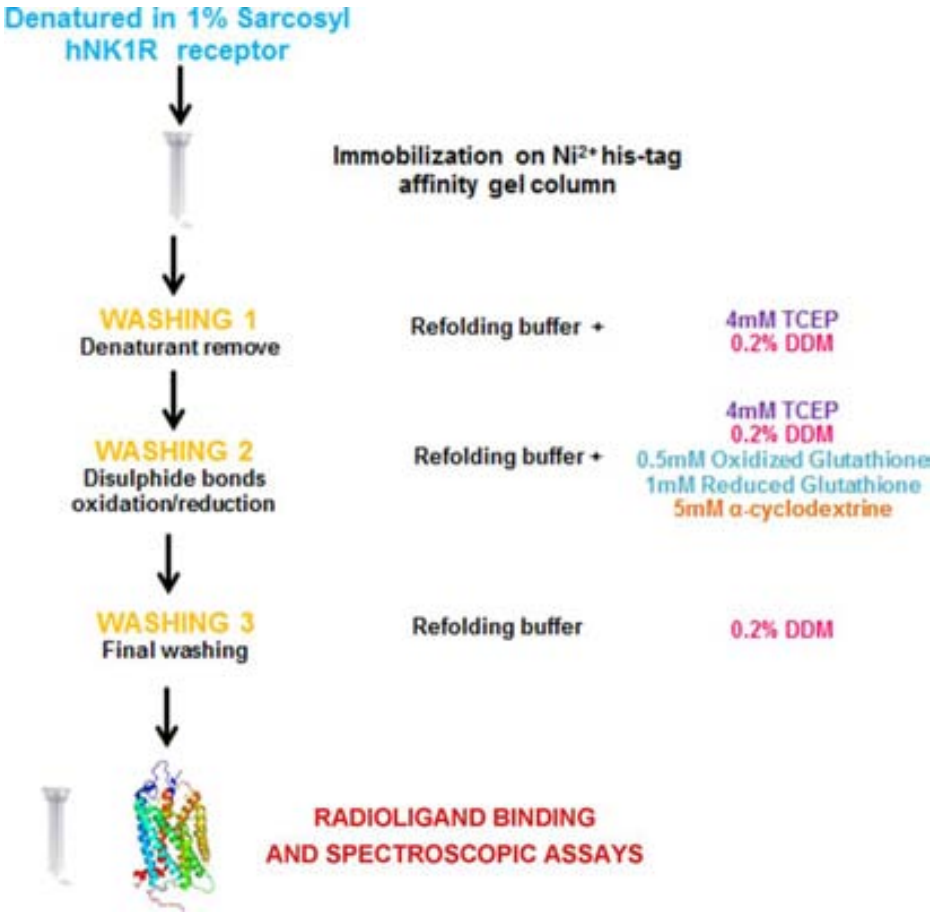


Figure 4.11 The on-column refolding process for hNK1R immobilized on His-select affinity gel. Three-buffers system is shown. All steps of the refolding process was carried out in 50 mM TrisHCl, 100 mM NaCl,10% glycerol (pH 7.4) buffer.

After screening for best refolding conditions, we established a protocol for on-column hNK1 receptor refolding. The three-buffers system (figure 4.11) was used in which the receptor refolding was mediated by 3 subsequent washes with the refolding buffer (50 mM TrisHCl, 100 mM NaCl, 10% glycerol, pH 7.4) containing different additives. As a material for refolding, we used hNK1R-366 and hNK1R-311 truncated forms of the receptor solubilized from *E. coli* inclusion bodies with 1% Sarkosyl. The receptor was immobilized on His-select nickel column via 6X-His tag on the C-end of the peptide and 4 mM TCEP was added to the samples to reduce all disulfide bonds. The first wash was done with the refolding buffer containing 0.2% DDM and 4 mM TCEP to remove impurities bound to the nickel column in a non-specific manner and to exchange the strong ionic Sarkosyl to mild non-ionic DDM detergent. The second wash buffer carried a glutathione oxidation-

reduction system required for proper formation of disulfide within the receptor transmembrane core. The *α*-cyclodextrine is often added to refolding buffers for complete elimination of a denaturant and helps the protein to adopt its native conformation. The final wash was done with the refolding buffer containing no additives except 0.2% DDM, essential for maintaining the receptor in its soluble state.

The receptor eluted from the column was subjected to spectroscopic studies and finally used in radioligand binding assay for its ability to bind substance P and was also characterized by SDS-PAGE after elution with the refolding buffer supplemented with 300 mM imidazole.

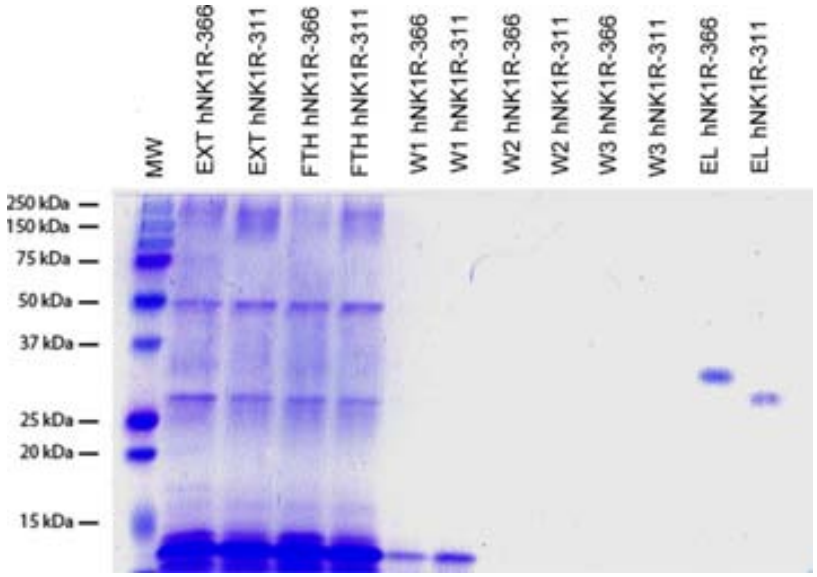


Figure 4.12 Purification and on-column refolding results for hNK1R-366 and hNK1-311 forms of the receptor. The samples are named as follows: *EXT*- whole inclusion bodies extract; *FTH* – flow-through of the column; *W1*, *W2* and *W3* – column washes; *EL* – protein eluted with 300 mM imidazole. 12% SDS-PAGE gel stained with stained with Coomassie solution.

Figure 4.12 presents purification and refolding results for both truncated forms of hNK1 receptor using three-buffers refolding system. The proteins eluted in 300 mM imidazole are similar to those purified in denaturing conditions, which means that 0.2% DDM concentration is sufficient to maintain hNK1R soluble in physiological buffer.

4.3.3. Estimation of hNK1R-366 and hNK1R-311 expression and purification outcome

Ultraviolet spectrophotometry remains a fast, accurate and nondestructive technique for protein concentration determination. The spectral properties of hNK1R-366 receptor arise from 9 tryptophan, 23 tyrosine and 22 phenylalanine residues. The hNK1-311 form of the receptor has 9 tryptophan 20 tyrosine and 18 phenylalanine residues. The absorbance spectra in UV region were recorded for both hNK1R-366 and hNK1-311 forms of the receptor purified on His-select nickel column in denaturing (1% SDS) and non-denaturing (1% DDM) conditions. The absorbance value at 280 nm was used to calculate the concentration of the eluted proteins to determine the yield of their expression and purification.

For both hNK1R-366 and hNK1R-311 the purification outcome is lower for refolded receptor than for denatured one, which can be explained by partial aggregation during the refolding steps influenced probably by detergent exchange.

Table 4.2 Final yield of hNK1R-366 and hNK1R-311 purification in denaturing and refolding conditions.

Sample	Yield (µg/L *E. coli*)
hNK1-366 in 1% SDS	720 ± 60
hNK1-366 in 0.2% DDM	460 ± 40
hNK1-311 in 1% SDS	313 ± 24
hNK1-311 in 0.2% DMM	150 ± 30

We should point out that the results obtained for both forms of the NK1 receptor purified in denaturing conditions show that the yield of hNK1R-366 in *E. coli* is 2.3 times higher than that of the hNK1R-311. When purified in refolding conditions hNK1R-366 gives better results yielding 3.5 times more purified product than hNK1R-311. We attributed this phenomenon to the fact that hNK1R-311 is less stable and is more prone to aggregate during purification process and detergent exchange in particular.

4.3.4. Circular dichroism spectra

Circular dichroism spectroscopy is one of the most general and basic tools to study protein folding. The absorption of light by different elements of the secondary structure acts as a measure of the degree of folding of the protein ensemble. In refolding studies this technique is usually used to measure the degree of unfolding of a protein by measuring the change in its absorption as a function of denaturant concentration. Far-UV CD spectra were recorded for purified hNK1-366 receptor in different concentrations of SDS. The receptor presents high helicity (about 60%) typical for GPCRs, which does not depend on the concentration of the denaturant when it is above CMC value (figure 4.13).

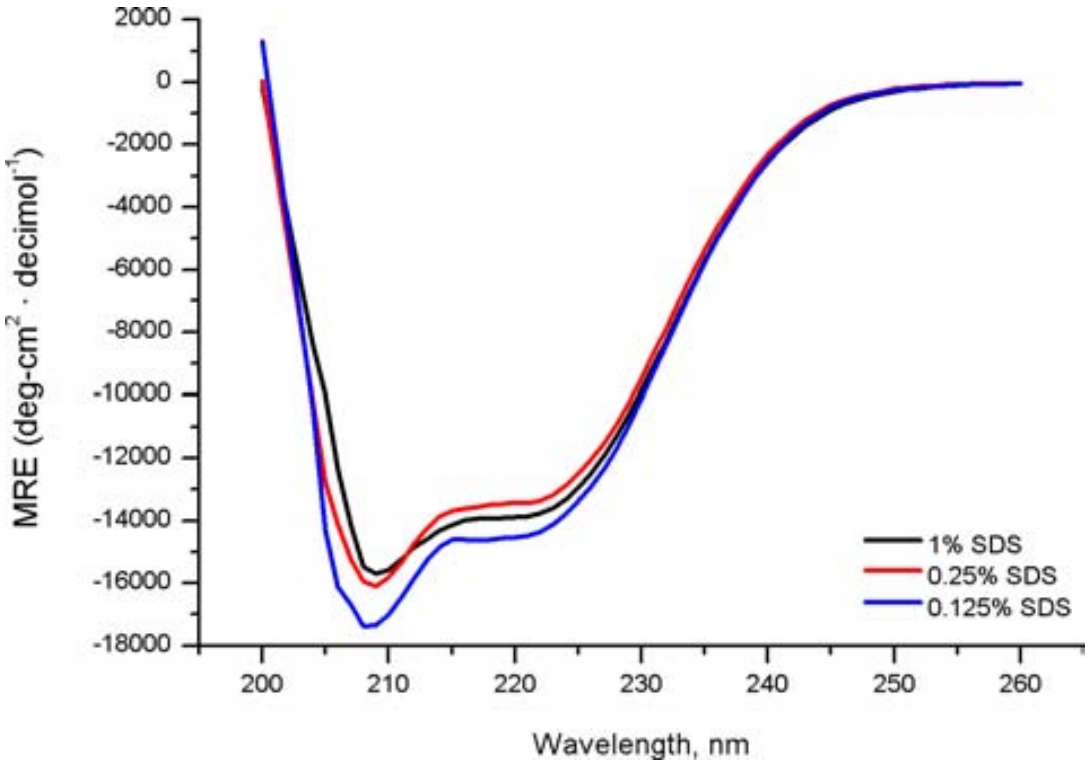


Figure 4.13 Circular dichroism spectra recorded for hNK1-366 in different concentration of SDS.

Sodium dodecyl sulfate earlier has been shown to induce secondary structure and *α*-helices in particular, in polypeptides, and is commonly used to model membrane and other hydrophobic environments. The precise mechanism by which SDS induces these conformational changes remains unclear until now.

However there are studies demonstrating that, the electrostatic interactions play a significant role in the formation and stabilization of SDS-induced structure (Montserret et al., 2000). Despite of conserving high helical content in SDS concentrations above CMC most of the proteins loose there functionality. As for example, upon solubilization in SDS, bacteriorhodopsin undergoes denaturation, preserving its secondary structure, but looses its retinal binding ability (London and Khorana, 1982).

The CD spectra recorded for hNK1R-366 in a mixture of SDS and DDM detergents (figure 4.14) show the decrease of MRE at 222 nm at higher DDM and lower SDS concentrations. As it can be seen from the table 4.3 in the mixture of 0.25% SDS and 0.75% DDM hNK1R-366 has about 60% of *α*-helix while at 1% DMM only 31% *α*-helical content can be detected.

Table 4.3 MRE at 222 nm and the percentage of *α*-helix for hNK1-366 and bovine opsin in different concentration of SDS and DDM.

Sample	SDS (%)	DDM (%)	MRE (deg-cm^2·decimol^{-1} at 222 nm)	*α*-helix (%)
hNK1R-366	1	—	-13 766	60
	0.25	—	-13 360	59
	0.125	—	-14 357	58
	0.25	0.75	-13 857	60
	0.125	1	-12 504	54
	—	1	-8 270	31
Bovine opsin	1	—	-9 000	36
	—	1	-8 000	31

Such an effect was observed for other proteins, as for example, porins – the *β*-barrel proteins that cross a cellular membrane and act as a pore through which molecules can diffuse. A SDS/DDM mixture maintains high concentrations of His6-porin in a *β*-strand-rich state. Higher DDM concentrations promote a higher level of *β*-strand (35%) with a concomitant decrease in *α*-helix.

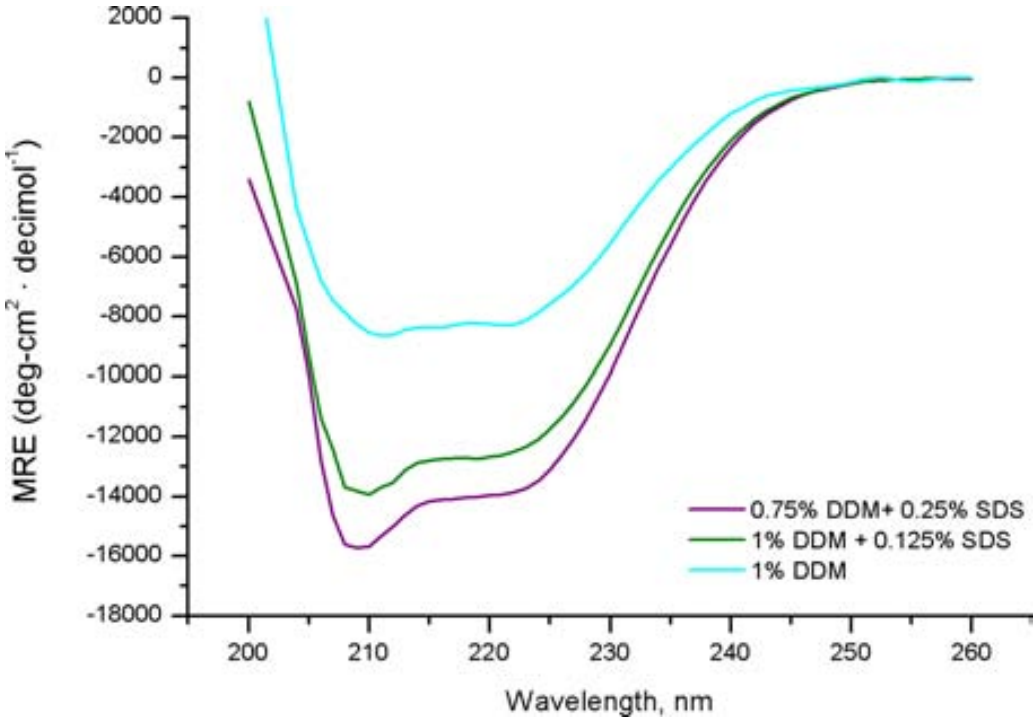


Figure 4.14 Circular dichroism spectra recorded for hNK1R-366 in different concentration of SDS and DDM.

Finally we compared the effect of SDS and DDM detergent on rhodopsin apoprotein (opsin) purified from bovine retinas. In 1% SDS opsin presents about 36% *α*-helical content which is lower than for hNK1R-366 purified in denaturing conditions (about 60%). At 1% DDM both bovine opsin and hNK1R-366 has 31% of *α*-helical structural elements (figure 4.15).

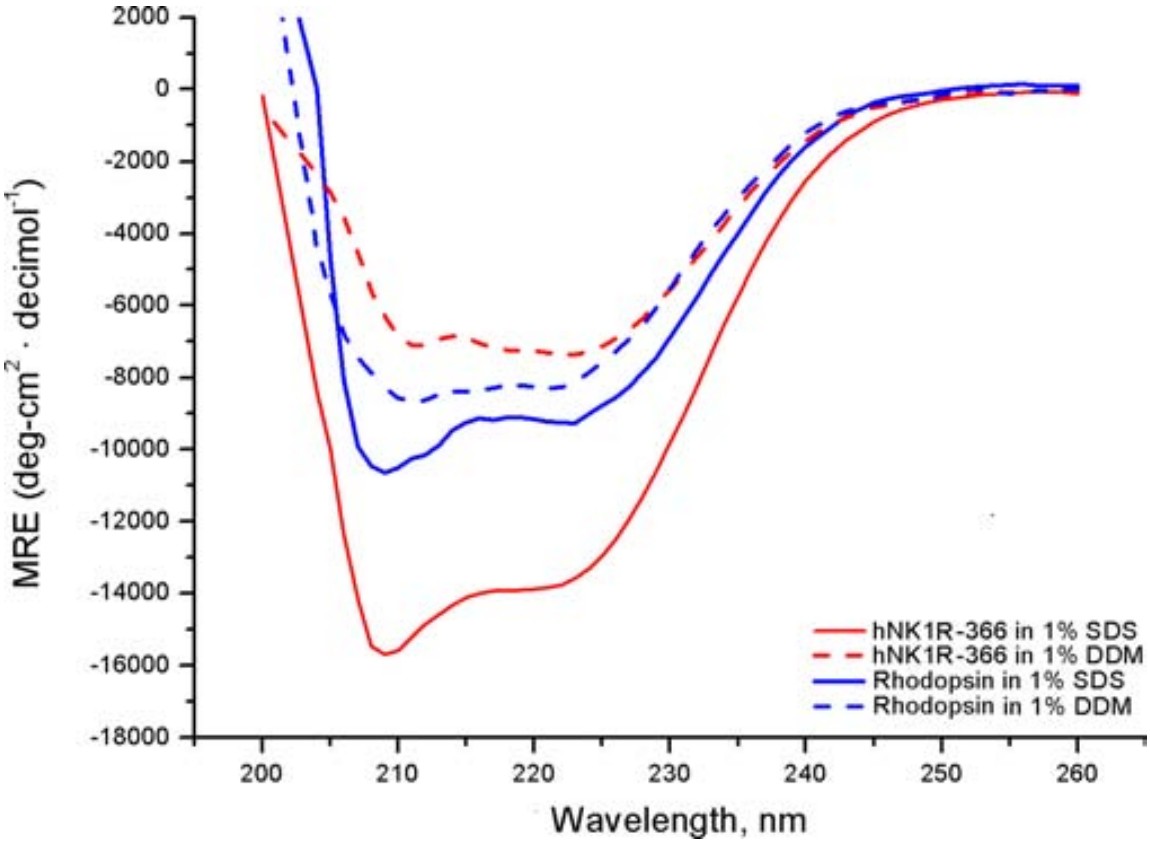


Figure 4.15 Circular dichroism spectra recorded for hNK1R-366 and rod outer segment (ROS) bovine rhodopsin in 1% SDS and DDM.

As a conclusion for this section we would like to point out that the results obtained from CD studies of denatured and refolded hNK1R-366 receptor show that the receptor refolded in mild non-ionic DDM detergent has amount of *α*-helical similar to that of the rhodopsin extracted from bovine retinas and solubilized in DDM micelles. This fact could be an indicator for correct hNK1R-366 folding.

4.3.5. Tryptophan intrinsic fluorescence studies on denatured and refolded hNK1 receptor

Intrinsic tryptophan fluorescence is a powerful techniques frequently used to monitor the folding/unfolding processes in proteins. The quantum yield of tryptophan is high enough to give a good fluorescence signal and 295 nm wavelength must be used to selectively excite tryptophans only. The main reason to use tryptophans to follow protein folding is that their fluorescence properties are sensitive to the environment which may undergo changes when protein folds or

unfolds. In a hydrophobic environment, Trp has higher fluorescence intensity while in a hydrophilic environment their quantum yield decreases leading to low fluorescence intensity. Strong emission maximum shifts as a function of the solvent can also be observed for Trp upon changes produced in their environment.

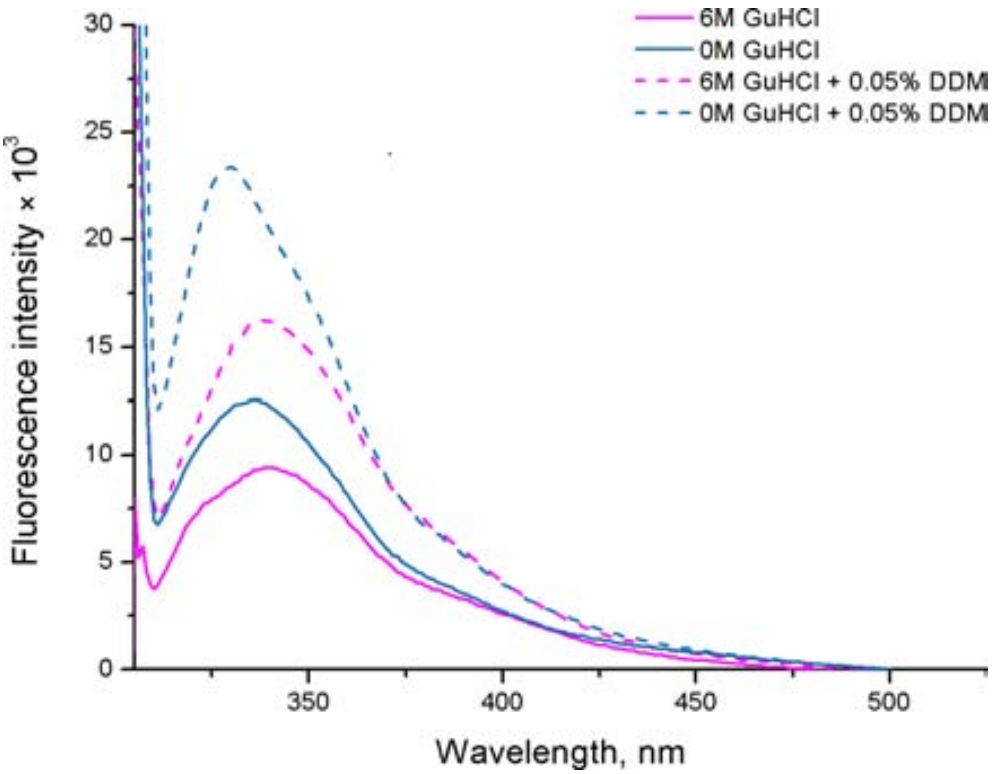


Figure 4.16 Intrinsic tryptophan fluorescence (λ_{exc}=295 nm) of hNK1R-366 at increasing concentrations of GuHCl in the presence and absence of DDM.

Tryptophan emission spectra (λ_{exc}=295 nm) were recorded for hNK1R-366 in decreasing concentrations of GuHCl in the presence and without DDM detergent. Table 4.4 shows the emission maxima wavelengths and fluorescence intensities for each sample presented in figure 4.16. The hNK1R-366 renatured in physiological buffer at 0 M GuHCl and without DDM exhibits strong emission maximum blue-shift (maximum at 336 nm) and significantly higher fluorescence intensity than the sample denatured in 6 M GuHCl (maximum at 340 nm).

Such a strong emission maxima blue-shift upon refolding was previously described for some membrane proteins. As for example, the refolded recombinant Toc75 chloroplastic outer envelope membrane protein showed emission peaks at

342 nm, while the same sample unfolded in urea emits with a decrease in intensity and a shift to 353 nm (Rogl et al., 1998). A similar decrease in intensity and a shift to longer wavelengths were also noted in the case of OmpA (Surrey and Jähnig, 1992). When as little as 0.05% DDM is added to the buffer during renaturation, hNK1R-366 exhibits stronger tryptophan emission blue-shift (figure 4.16 and table 4.4) than in physiological buffer only.

Table 4.4 Emission maximum and fluorescence intensity values for hNK1R-366 in 6 M and 0 M GuHCl in the presence and in absence of 0.05% DDM.

Sample	Emission maximum (nm)	Fluorescence intensity $\times 10^{-3}$
6 M GuHCl	340	9
0 M GuHCl	335	12
6 M GuHCl + 0.05% DDM	339	16
0 M GuHCl + 0.05% DDM	329	23

Despite of the blue-shift characteristic for tryptophans in hydrophobic environment, commonly observed upon protein renaturation, the hNK1R-366 refolded in physiological buffer was found to be relatively unstable and prone to aggregate in about 24 hours (figure 4.17 *B*). The changes in turbidity of the samples were monitored at 400 nm by UV-vis spectroscopy of. At low denaturant concentrations, hNK1R starts to aggregate which results in an increase of absorbance at 400 nm. However, the presence of 0.05% DDM stabilizes the receptor that can be deduced from the absence of absorbance at 400 nm.

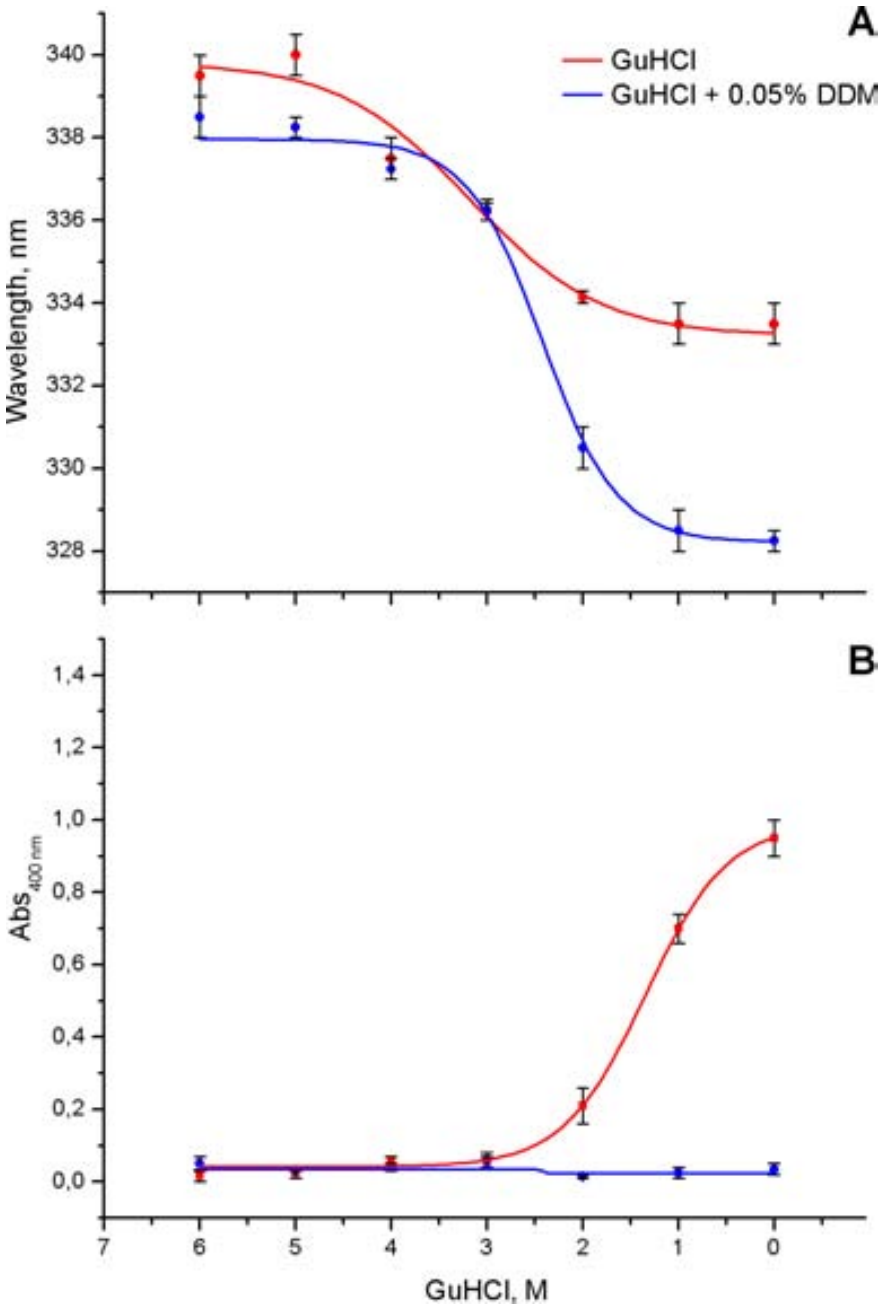


Figure 4.17 (*A)* Shift of fluorescence emission maximum of the hNK1R-366 tryptophans as a function of GuHCl concentration. (*B)* Changes in the sample turbidity (absorbance at 400 nm) with the decrease of GuHCl concentration, in presence of 0.05% DDM or without DDM.

As a final stage of intrinsic tryptophan fluorescence studies we compared the hNK1R-366 in different concentrations of GuHCl and hNK1R-366 expressed in the membrane of COS-1 cells and solubilized in 1% DDM. As it can be seen in figure 4.18 the fluorescence emission maximum of the receptor undergoes a strong

blue-shift upon denaturant concentration decrease. At 0 M GuHCl the refolded hNK1R-366 exhibits the same emission wavelength maximum around 335 nm as the receptor expressed in the membrane of COS-1 cells and solubilized in 1% DDM.

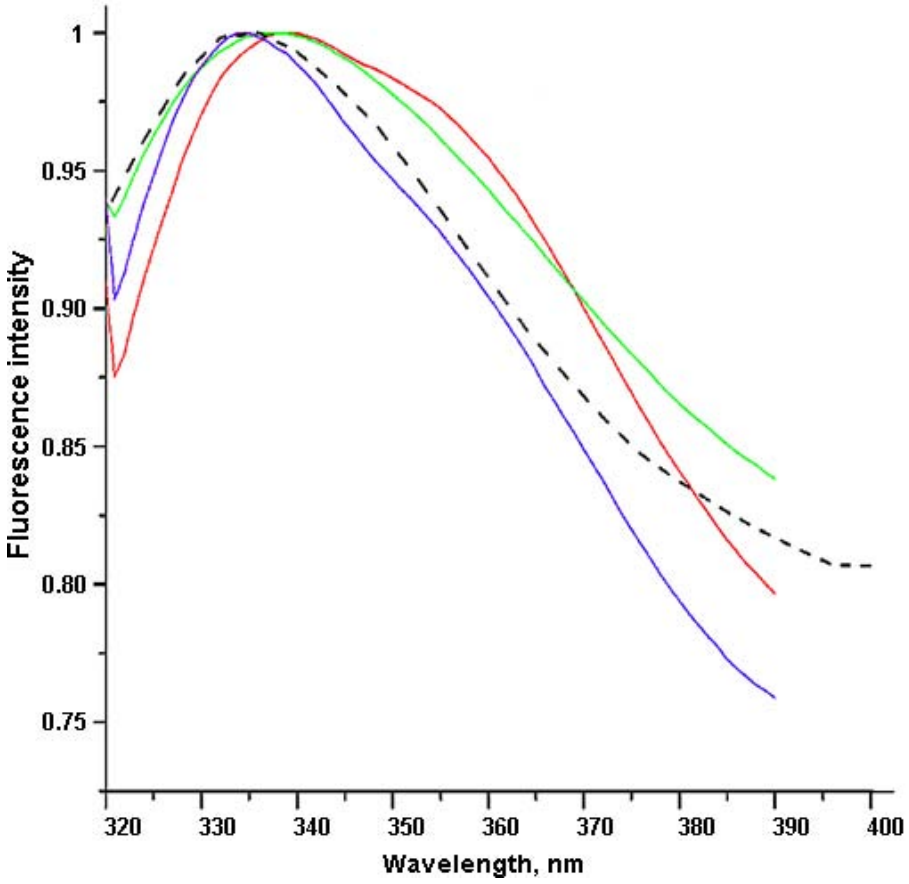


Figure 4.18 Changes in maximum of Trp emission of hNK1R-366 (λ_{ex}=295 nm) dependent on GuHCl concentration. hNK1R-366 at 6 M GuHCl – *red*; 2 M GuHCl – *green*; 0 M GuHCl – *blue*; recombinant hNK1R expressed in COS-1 cells and solubilized in 1% DDM – *dotted line*.

The results of intrinsic tryptophan fluorescence studies show that in low concentrations of GuHCl there is a blue-shift in the emission spectrum peak of hNK1R-366 which indicates changes of tryptophan to more hydrophobic environments as a result of protein folding (figure 4.17 *A*, *red trace*). The emission spectra of hNK1R-366 expressed in the membrane of COS-1 mammalian cells shows very similar emission maximum to that of the receptor refolded from the inclusion bodies. The shift becomes more evident if 0.05% DDM is added to the refolding buffer, indicating that the micellar structure of this non-ionic detergent stabilize the refolded receptor structure (figure 4.17 *A*, *blue trace*).

4.3.6. Saturation radioligand binding assays

The interaction between ligands and receptors can be quantified by radioligand binding assay. Two important parameters should be considered. The first one is the K_d which is the affinity of a ligand for a receptor and is the concentration of ligand that will occupy 50% of the receptors. The second parameter is the B_{max} value that shows the maximum density of functional receptors.

In the present work we performed saturation radioligand binding for refolded hNK1R-366 and hNK1R-311 truncated forms of the receptor in order to determine their ability to bind ^{3}H-labeled Substance P and thus the quality of the refolding process. As a radioligand we used Substance P (9-Sar, 11-Met(O2)), [2-Prolyl-3,4-^{3}H] (abbreviated [^{3}H]-SP) with specific activity of 41 Ci/mmol that was previously used in studies related to Neurokinin-1 receptor. (Tousignant et al., 1991; Coge and Regoli, 1994).

4.3.6.1. Radioligand binding of [^{3}H]-SP for full-length hNK1R expressed in COS-1 cells

We used full-length hNK1R expressed in mammalian COS-1 cell line as a reference in radioligand binding assay for refolded forms of the receptor. According to published results from [^{3}H]-SP saturation binding studies, rat NK1 receptor expressed in COS-1 cells showed a the K_d of 0.56 ± 0.18 nM with B_{max} of 1.2 ± 0.2 pmol/mg of protein (Pradier et al., 1994).

Figure 4.19 presents total and non-specific binding curves of the full-length hNK1R in intact COS-1 cells with increasing concentrations of [^{3}H]-SP. The cells were incubated in the presence of increasing concentrations of the ligand ranging from 0.1 to 6 nM. Non-labeled SP was used to determine the non-specific binding.

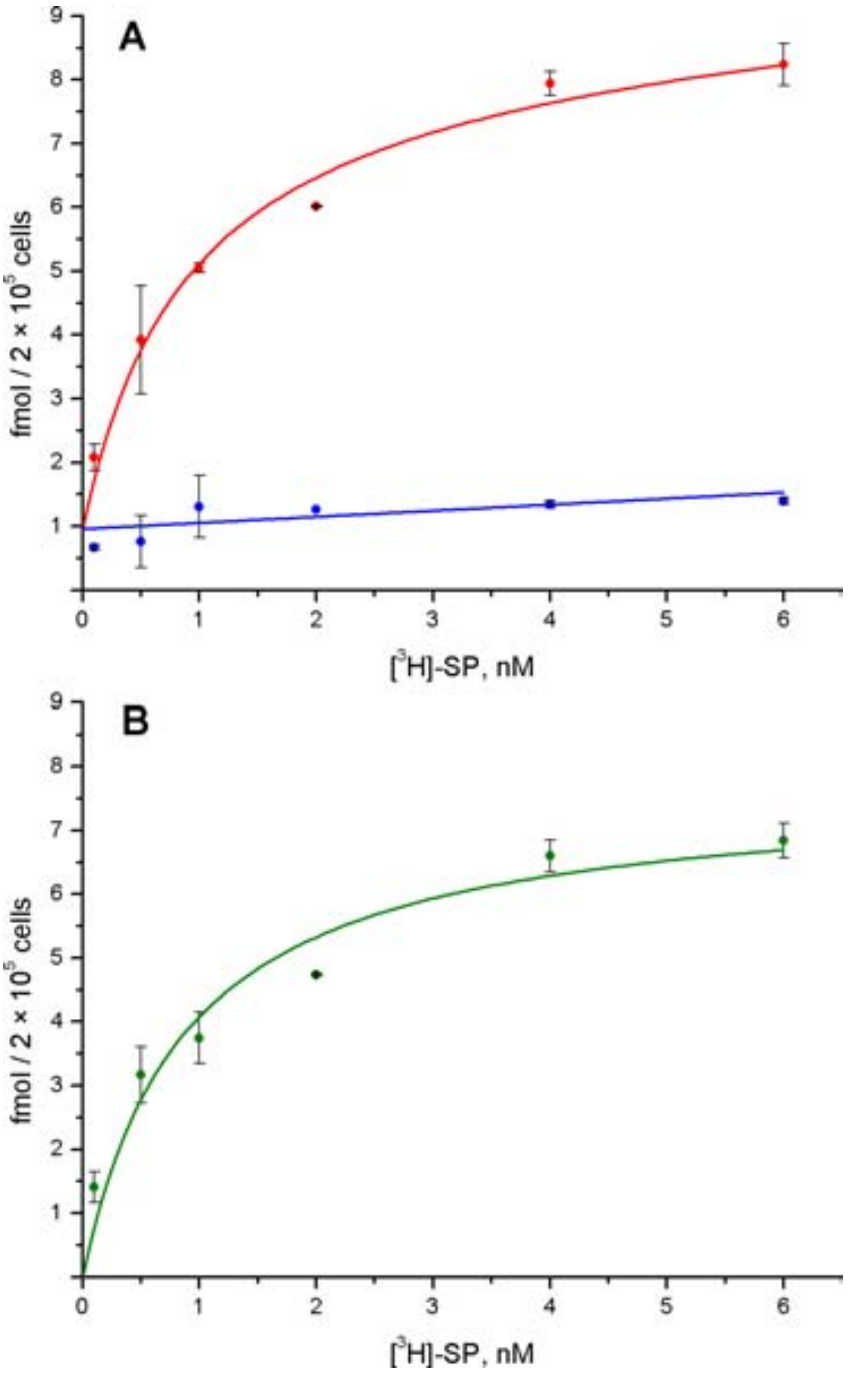


Figure 4.19 Saturation radioligand binding assay of full-length hNK1R expressed in membranes of COS-1 cells. *(A)* Total and non-specific binding curves. *(B)* Specific binding curve.

From a numerical fit of the data points of specific binding saturation curve (figure 4.19) the K_d was determined as 0.88 ± 0.26 nM for [^{3}H]-SP.

4.3.6.2. Radioligand binding of [^{3}H]-SP for refolded hNK1R-366

To assess the functionality of the refolded hNK1R-366 and hNK1R-311 refolded forms of the receptor we performed radioligand binding saturation assay in increasing concentrations of [^{3}H]-SP.

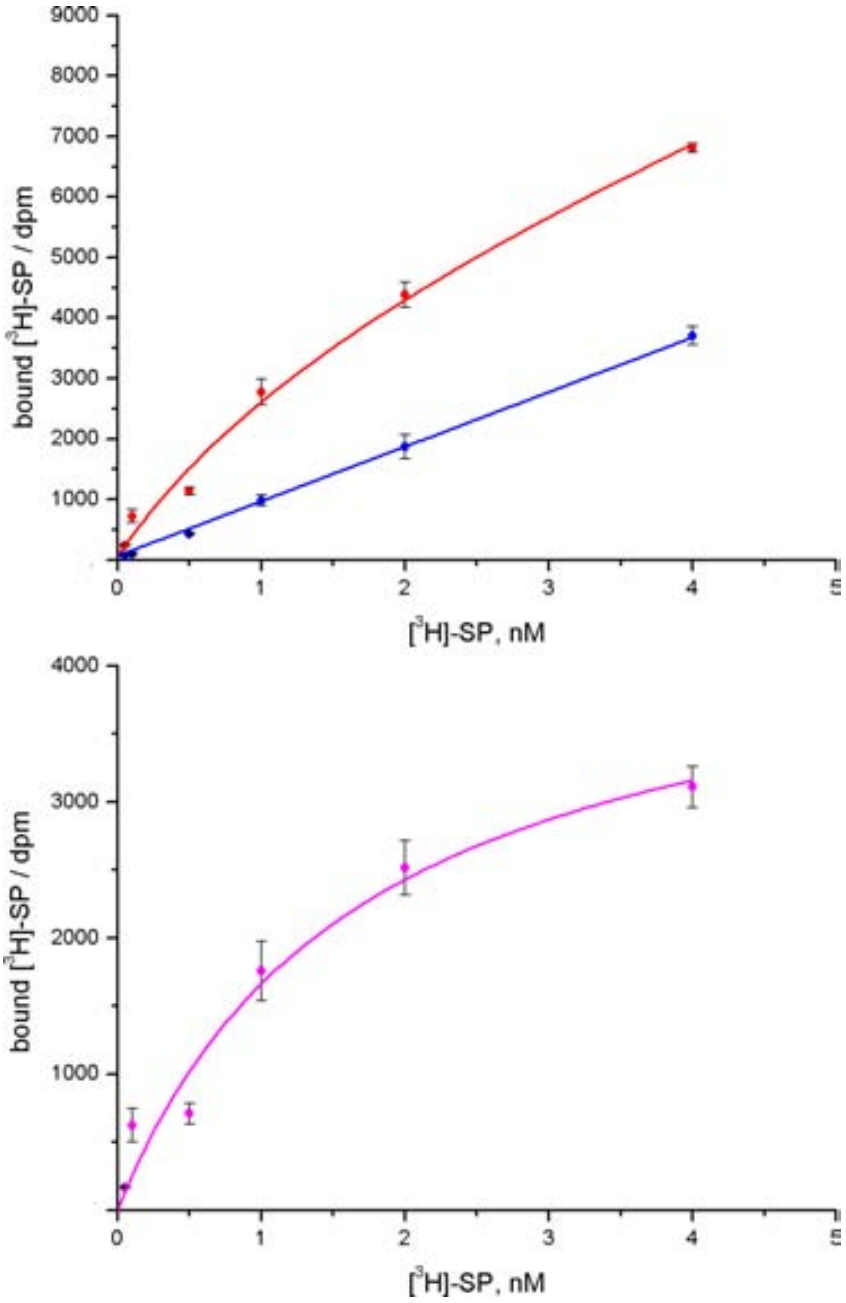


Figure 4.20 Saturation radioligand binding assay for truncated hNK1R-366 refolded from *E. coli* inclusion bodies. *(A)* Total and non-specific binding curves. *(B)* Specific binding curve.

A representative saturation curve of the radioactivity of the specifically bound [^{3}H]-SP as a function of [^{3}H]-SP concentration is shown in figure 4.20. The non-linear curve fitting with an equation based on the law of mass actions with a 1:1 stoichiometry of receptor and ligand allows calculating the active receptor concentration. The calculated K_d for the refolded hNK1R-366 equals to 1.7 ± 0.6 nM. This value is about 2 times higher than for the recombinant receptor expressed in the membrane of COS-1 cells. Knowing the total amount of the refolded receptor subjected to ligand binding studies and the B_{max} value we estimated the active receptor concentration. In our experiments the B_{max} value was 0.25 ± 0.04 nM. This value is lower than the expected one. This result will be discussed in the next section.

4.4. hNK1R C-terminus characterization

According to the objectives of the present study we aim to characterize the hNK1R C –terminus using different spectroscopic techniques as UV, fluorescence, FTIR and circular dichroism spectroscopies to elucidate its structural properties. Methods of secondary and tertiary structure prediction along with the molecular modeling became helpful for better understanding and visualization of hNK1R C-terminus overall structure.

4.4.1. Purification of hNK1R C-terminus solubilized from *E. coli* inclusion bodies

The solubilized peptide in 6 M GuHCl was applied to the nickel affinity column at gravity flow rate for tight binding to occur. Extensive washings with about 20 volumes of Ni^{2+} gel were necessary to remove proteins bound to the column in non-specific manner and to remove traces of GuHCl. The purified hNK1R C-terminus was eluted from the column using 100 mM imidazole or alternatively with 100 mM L-histidine in non-denaturing conditions.

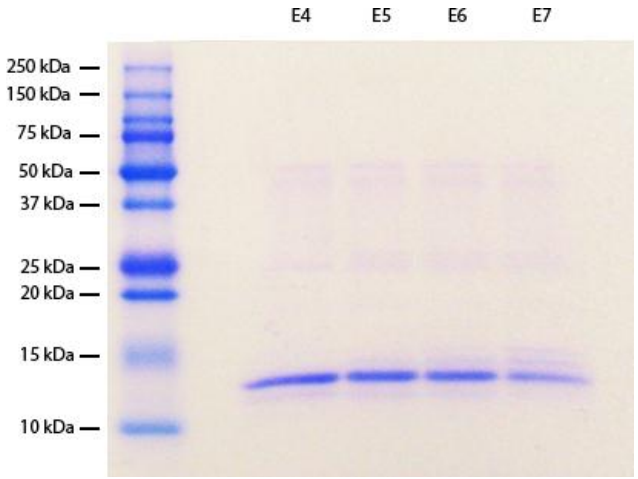


Figure 4.21 Analysis of IMAC purification of hNK1R C-terminus 6X-His tag. Four fractions of ten (from E4 to E7) done with 100 mM imidazole in 20mM phosphate (pH 7.4), 100 mM NaCl, 4mM TCEP buffer are shown on SDS-PAGE 18% gel stained with Coomassie solution. The 13 kDa band corresponds to the hNK1R C-terminus monomer.

Figure 4.21 shows purification results for hNK1R C-terminus. Ten elution fractions (1 ml each one) were collected. When 4mM TCEP is added at every stage of peptide's extraction and purification process to reduce disulfide bonds only a monomeric band at 13 kDa can be observed. In absence of TCEP higher molecular weight bands can be observed on SDS-PAGE gel. This can be attributed to oligomers (tetramers, trimers and dimers) of hNK1R C-terminus possibly formed due to intermolecular disulfide bonds formation (figure 4.22).

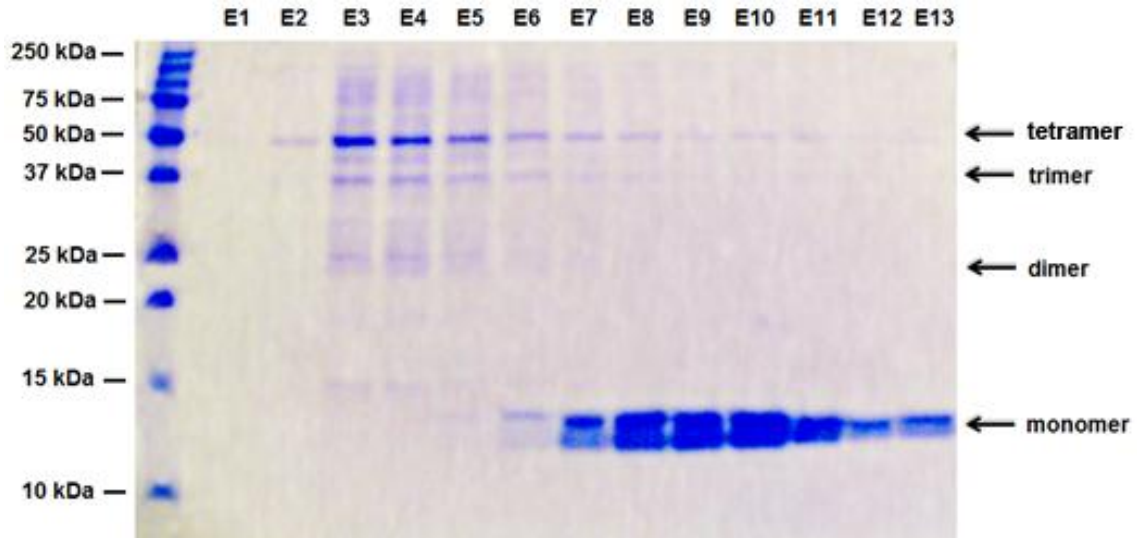


Figure 4.22 Analysis of IMAC purification of hNK1R C-terminus. Elution fractions 1 ml each one in 20 mM phosphate (pH 7.4), 100 mM NaCl, 100 mM histidine buffer are shown on SDS-PAGE 18% gel stained with Coomassie solution. The hNK1R C-terminus oligomers (tetramers, trimers) and dimers are eluted in the first fraction. Late fractions contain high purity hNK1R C-terminus monomer.

As it can be noticed from the electrophoresis results obtained for hNK1R C-terminus purified in absence of TCEP, the first five fractions only contain high molecular weight bands that are attributed to oligomers (tetramers, trimers and dimers) of hNK1R C-terminus possibly formed due to intermolecular disulfide bonds formation. The posterior five fractions (E7-E11) present the monomeric band at 12 kDa and a smaller band at 11 kDa possibly attributed to products with an intermolecular disulfide bond formed that has greater mobility in SDS-PAGE gel. Fractions from E12 to E13 contain mostly monomeric form of hNK1R C-terminus.

4.4.2. UV spectra of hNK1 C-terminus

The hNK1R C-terminus is a peptide of 96 amino acid residues. Its spectroscopic properties arise from 6 phenylalanine and 3 tyrosine residues present in its sequence. The phenylalanine absorption is too low in comparison with the tyrosine one and cannot be analyzed. Tyrosine residues are not as much sensitive to the dissolvent polarity as tryptophans; however changes in their absorption bands can be still used to obtain information about their environment. Figure 4.23 shows raw and deconvolved absorption spectra of hNK1R C-terminus.

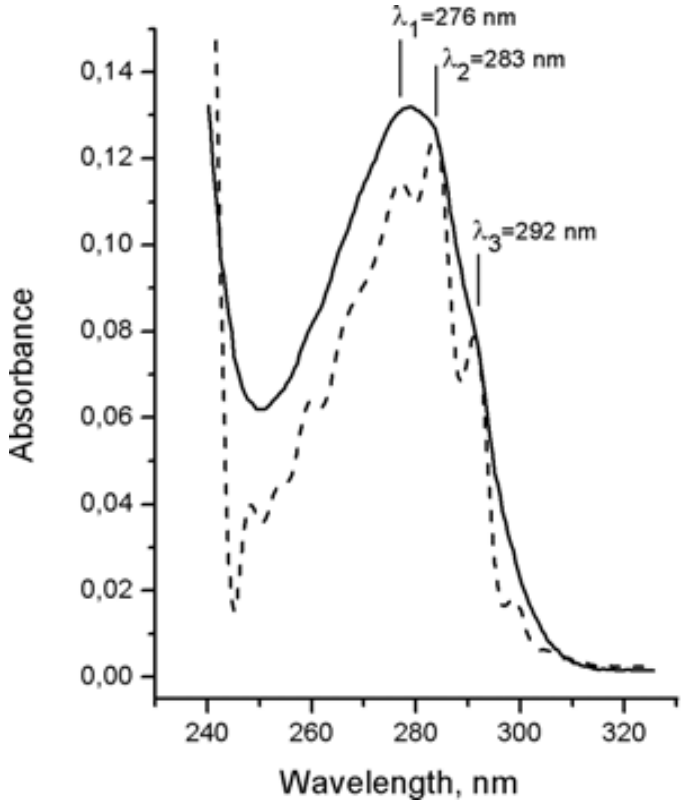


Figure 4.23 Absorption spectrum *(solid line)* and deconvolved spectrum *(dotted line)* of hNK1 C-terminus in 25 mM sodium phosphate, 100 mM NaCl (pH=7.4) buffer. Three well-defined peaks can be distinguished on the deconvolved curve, at 276, 282 and 292 nm.

The λ_1 maximum is located at 276 nm and probably originates from the unique vibrational band. The λ_2 maximum at 283 nm mainly corresponds to the O-O vibrations of the phenolic ring. Based on previous spectroscopic studies made on *N*-acetyl-L-tyrosine ethyl ester in different solvents (Padrós et al., 1982) it can be concluded that λ_1 and λ_2 maxima of hNK1R C-terminus are characteristic of

tyrosine in non-polar environment. The third λ_3 wavelength maximum at 292 nm could be attributed to the ionized state of –OH group of the phenolic ring of tyrosine. At pH<10 the hydroxyl group of tyrosine is normally protonated and no presence of tyrosinate should be observed. However, at physiological pH anomalous tyrosine absorbance was detected for some proteins at 290 nm. Such spectral perturbations were assigned to hydrogen-bonded complexes with proton acceptor residues, as for example, glutamic or aspartic residues side chains (Szabo et al., 1979). At low pH values the ionized form of tyrosine should disappear, whereas at pH higher than the pK of tyrosine (pK=10.0), predominantly tyrosinate absorption should be observed. Unfortunately, we were unable to perform tyrosine ionization assay at different pH values, because hNK1R C-terminus was prone to aggregate at pH below 6 and over 9. However, when working in this range of pH values no difference in absorbance spectra was detected.

4.4.3. CD spectra of native and denatured hNK1R C-terminus

Far-UV circular dichroism spectra of NK1R C-terminus were recorded in physiological buffer and denaturing conditions at different concentration of GuHCl in 190-260 nm region. Figure 4.24 shows two curves that correspond to the native state and 6 M GuHCl denatured state of the peptide. As it can be clearly seen the CD spectra of the native peptide has two negative maximum at 222 nm and 208 nm characteristic of *α*-helical secondary structure (Greenfield, 2006).

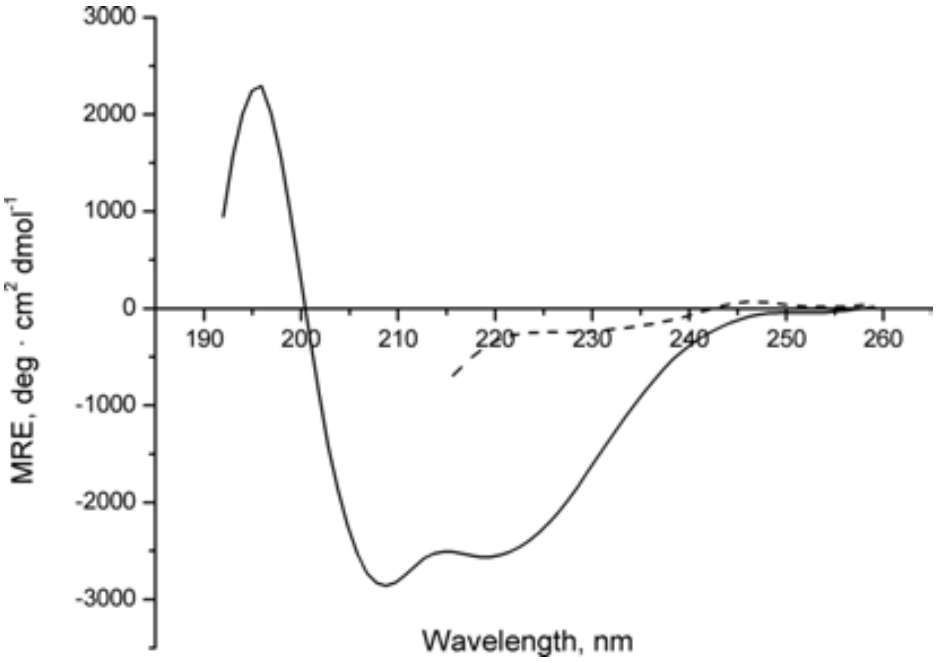


Figure 4.24 Far-UV CD spectra recorded for native (*solid line*) and 6 M GuHCl denatured (*dashed line*) hNK1R C-terminus. Only 215 nm to 260 nm scan range is shown for denatured peptide due to high absorbance of GuHCl in far-UV region.

The estimation of percentage of secondary structure elements was performed using SELCON3 program based on 37 proteins reference set (Sreerama and Woody, 1993). The hNK1R C-terminus shows about 23% helical content, 35% of unordered structure and 42% of other structural components as β-sheets. A loss of tertiary and secondary structure of proteins may be induced by increasing concentrations of GuHCl, known to be a chaotropic agent widely used for proteins denaturation. Usually high concentrations (up to 6 M) are needed to completely disrupt the secondary structure. CD spectra were recorded for hNK1R C-terminus at different concentrations of GuHCl to evaluate the loss of its secondary structure. Mean residue ellipticity at 222 nm (θ_{222}) was used as a measure of ellipticity.

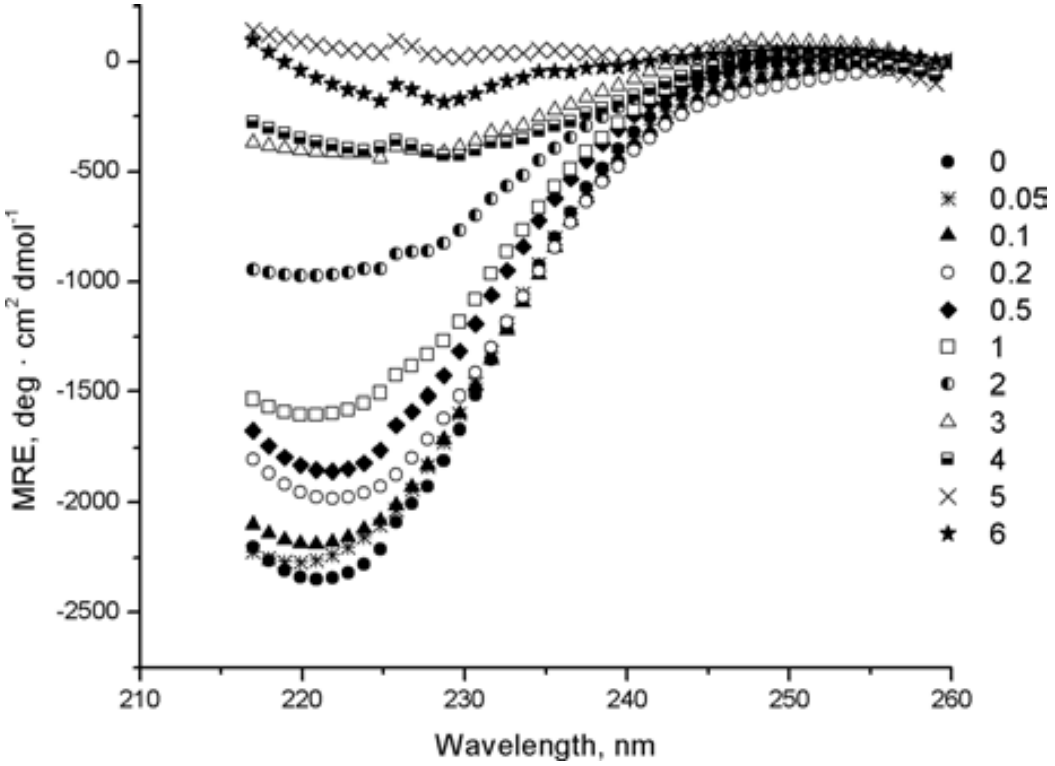


Figure 4.25 Far-UV CD spectra recorded for native hNK1R C-terminus at increasing concentrations of GuHCl ranging from 0 M to 6 M. Only 215 nm to 260 nm scan range is shown for denatured peptide due to high absorbance of GuHCl in far-UV region.

As it can be seen in figure 4.25 a considerable reduction of peptide's regular secondary structure conformation occurs in the presence of GuHCl. A significant loss of helicity already starts at GuHCl concentration as low as 0.2 M. At 1 M GuHCl the helical content has decreased significantly and at 6 M GuHCl the hNK1R C-terminus completely loses its *α*-helical component. The obtained CD spectra both in native and denaturing conditions suggest that hNK1 C-terminus has clearly defined secondary structure elements as *α*-helix and *β*-sheets which according to SELCON3 method constitute about 65% of overall structure of the peptide.

4.4.4. Intrinsic tyrosine fluorescence studies

The hNK1R C-terminus has three tyrosine residues which can be used as intrinsic sensor upon tertiary structure changes. Figure 4.26 *(top trace)* shows the emission spectrum of the peptide in buffer. The emission maximum is found at 328 nm. It cannot be attributed to the fluorescence of the protonated form of

tyrosine, since its emission maximum is located around 308 nm. In comparison with free tyrosine at pH 7.4, the hNK1R C-terminus emission spectrum is much broader in the long-wavelength region and could possibly arise from tyrosinate fluorescence (Szabo et al., 1979).

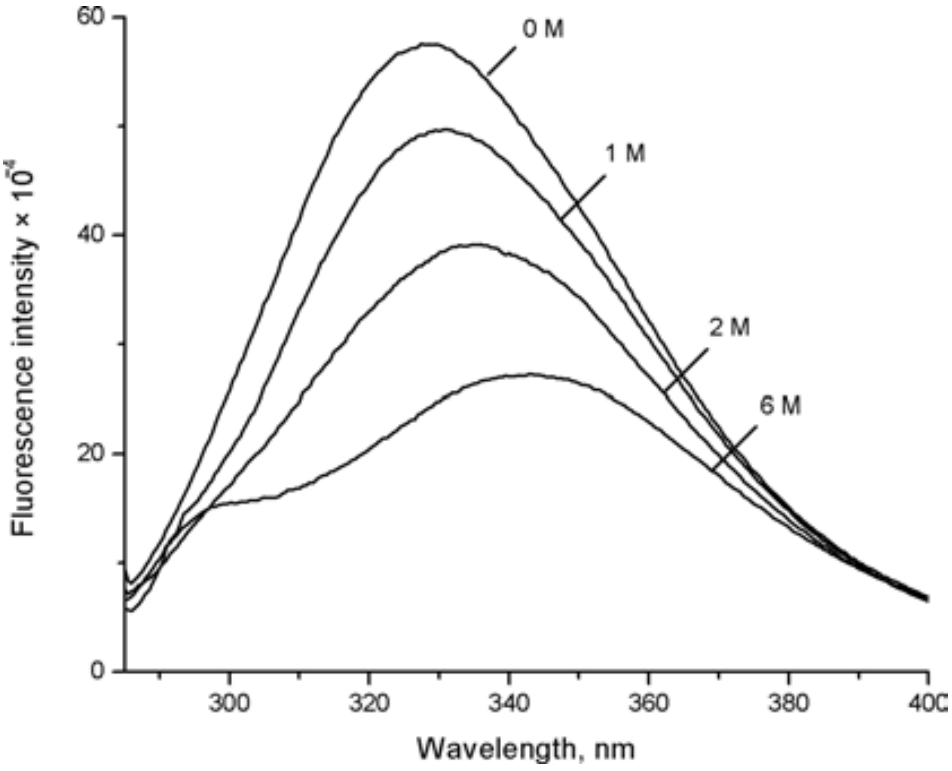


Figure 4.26 Intrinsic tyrosine fluorescence of hNK1R C-terminus (λ_{exc}=280 nm) in different concentrations of GuHCl ranging from 0 M to 6 M (4 points of denaturation curve are shown) in 25 mM sodium phosphate, 100mM NaCl (pH 7.4), 1mM TCEP buffer.

After spectra deconvolution (figure 4.27 and table 4.5), 3 bands at 296, 328 and 347 nm can be detected on the spectra. We assigned the band at 296 nm to normal tyrosine emission naturally found in proteins. The band at 347 nm possibly arises from tyrosinate emission in polar environment and the band at 328 nm can be assigned to a tyrosinate found in a non-polar microenvironment. The pK_a of phenolic ring is too high (p*K*a=10.0) to suggest that the anomalous tyrosine emission could be originated from the ground state tyrosinate which requires basic environment to be formed around phenyl moiety.

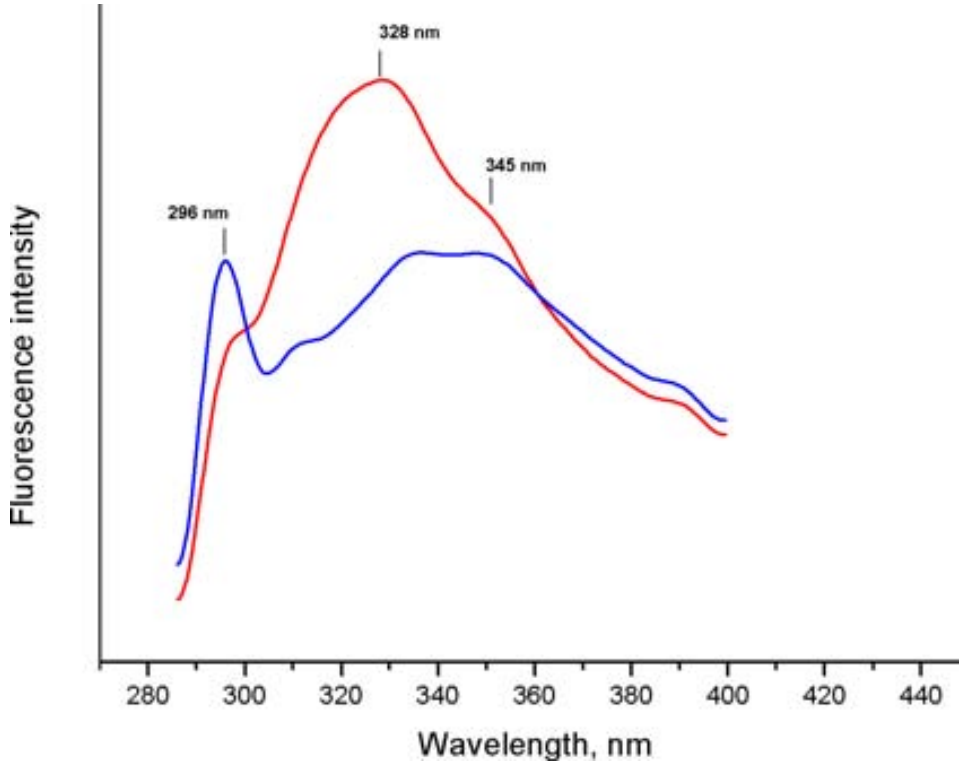


Figure 4.27 Deconvolved spectra of hNK1R C-terminus in absence of GuHCl (*red*) and denatured in 6 M GuHCl (*blue*). The spectra were deconvolved using the following parameters: k-factor = 1.200 and HW = 2.100.

Anomalous red-shifted tyrosine fluorescence at 345 nm at neutral pH was first reported in two toxic peptides isolated from the Indian Cobra (*Naja naja*) (Szabo et al., 1979). Some other tyrosine-containing proteins with no tryptophan residues were also candidates for anomalous tyrosine emission, as for example, pig intestinal Ca^{2+}-binding protein (Prendergast et al., 1984), *α*- and *β*- purothionins from wheat (O'Neil and Hofmann, 1987) and histones of H1 class (Jordano et al., 1983). The abnormal tyrosine fluorescence emission in tryptophan-lacking peptides or proteins is often explained as originated from an intermolecular excited state proton transfer (Szabo et al., 1979) from the phenolic moiety hydroxyl to a proton acceptor such as carboxyl groups of glutamic or aspartic residues (O'Neil and Hofmann, 1987).

Table 4.5 Assignment of emission bands of hNK1R C-terminus in 0 M and 6 M GuHCl to different tyrosine species.

Sample	Emission range (nm)	Assignment
hNK1R C-terminus without GuHCl	296	Normal tyrosine fluorescence
	328	Band originated from hydrogen-bonded species of tyrosine or an exciplex formation
	347	Tyrosinate emission
hNK1R C-terminus in 6 M GuHCl	296	Normal tyrosine fluorescence
	334	Band originated from hydrogen-bonded species of tyrosine or an exciplex formation
	347	Tyrosinate emission

The band at 328 nm was earlier reported for plastocyanin (Graziani et al., 1974) or an excited state tyrosine complex formation with an unidentified group. This phenomenon was already described for several proteins where the phenolic moiety of the tyrosyl residue appeared to be hydrogen bonded but tyrosinate emission was not confirmed. As for example, adrenodoxin, a single tyrosine protein, showed anomalous emission at 331 nm, which was interpreted as an exciplex (excited state complex) between tyrosine and an unknown partner (Lim and Kimura, 1979). The anomalous tyrosine UV absorption and emission of hNK1 C-terminus can be correlated with the primary amino acid sequence of the peptide rich in aspartic and glutamic residues which are known to be good proton acceptors due to conjugate bases of their carboxyl groups.

Two regions should be pointed out: the first one starts from Asp^{330} and extends up to Glu^{335} enclosing Tyr^{331} residue; the other one extends from Asp^{366} to Glu^{371}. These two regions rich in proton acceptor residues being in proximity to tyrosyl moieties may cause perturbations in fluorescence emission spectra maxima.

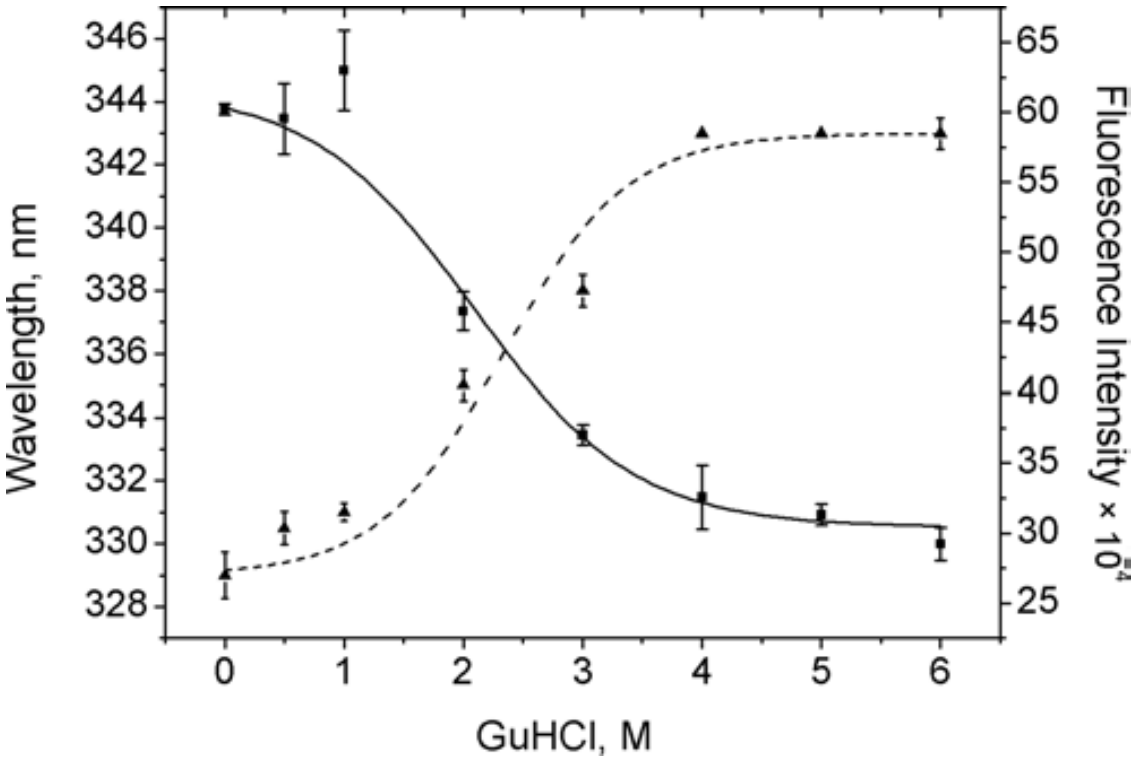


Figure 4.28 Titration curves of hNK1R C-terminus with GuHCl. Emission maximum position change (*dashed line*) and emission fluorescence intensity at 328 nm (λ_{exc}=280 nm) (*solid line*) are shown in different concentrations of GuHCl ranging from 0 M to 6 M. All spectra were recorded in 25 mM sodium phosphate, 100mM NaCl (pH 7.4), 1 mM TCEP buffer.

Upon increasing GuHCl concentration the appearance of a weak band at 300 nm can be observed on the spectra, especially at 6 M GuHCl. This band could be attributed to normal tyrosine fluorescence. The *dashed line* in figure 4.28 presents emission maximum shift for hNK1R C-terminus in different concentrations of GuHCl. The emission maximum changes towards longer wavelength upon unfolding and reaches its highest value of 343 nm at 6 M GuHCl. The transition point is located around 2.3 M GuHCl. At 4 M of the denaturing agent no more wavelength shift can be observed which is indicative of the complete loss of peptide's tertiary structure. The quantum yield of tyrosine fluorescence is very sensitive to interactions of tyrosine with its environment. The fluorescence emission intensity of hNK1R C-terminus at 328 nm in increasing concentration of GuHCl is represented by the *solid line* in figure 4.28. The transition point is observed at around 2.1 M GuHCl and at 4 M the emission intensity arrives to a plateau, in agreement with the plot of the maximum wavelength.

4.4.5. FTIR analysis of secondary structure

FTIR spectra were recorded for hNK1R C-terminus to determine secondary structure elements and their contribution into the overall structure. The hNK1R C-terminus has 3 cysteines within its sequence which can be easily oxidized yielding oligomeric forms due to disulfide bonds formation. Reduction agents as TCEP or DTT can be used to maintain reduced cysteines, however the use of these compounds is limited since their absorption range overlays with that of the amide I and amide II regions. *N*-ethylmaleimide is a good alternative as a cysteine blocking agent which can be used when aggregation and oligomeric forms formation is are not desirable for as they can cause misinterpretation of structural properties of a studied protein.

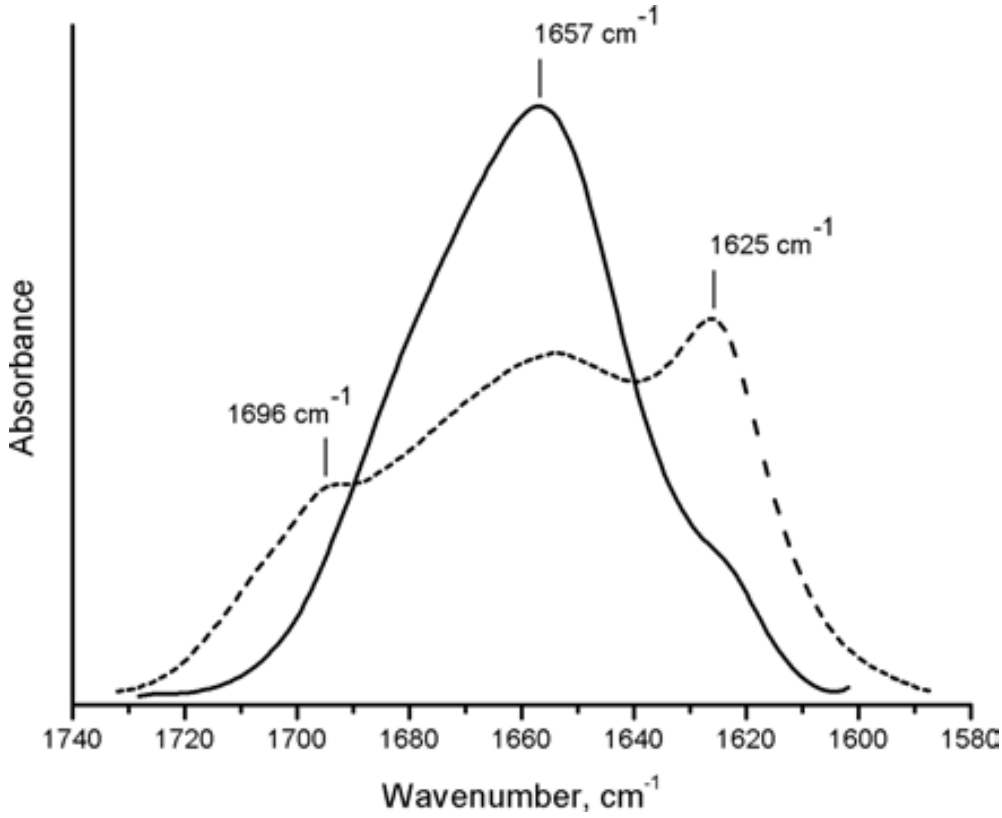


Figure 4.29 Baseline corrected FTIR spectra of hNK1R C-terminus in 20 mM sodium phosphate buffer (pH 7.4) (solid line), and aggregated hNK1R C-terminus (*dashed line*). The amide I band is shown.

In this regard all hNK1R C-terminus samples used in FTIR experiments were treated with NEM 1 mg/mg of IBs prior to purification to prevent protein aggregation and higher molecular weight oligomers formation. Figure 4.29 shows FTIR absorption spectrum of hNK1R C-terminus in 25 mM sodium phosphate buffer (pH 7.4) and the spectrum of the same protein aggregated within 24 hours. The band between 1600 and 1700 cm^{-1} encloses the amide I region that is mainly associated with C=O stretching vibrations directly related to the backbone conformation and is used to interpret the secondary structure composition of the peptide. The absorption maximum is located at 1657 cm^{-1} which is indicative of a predominant *α*-helical content of hNK1R C-terminus. The aggregated form of the peptide presents two well-defined bands with maxima at 1696 and 1625 cm^{-1} which are characteristic of extended antiparallel *β*-structures associated with denatured or aggregated proteins (Szabó et al., 1999).

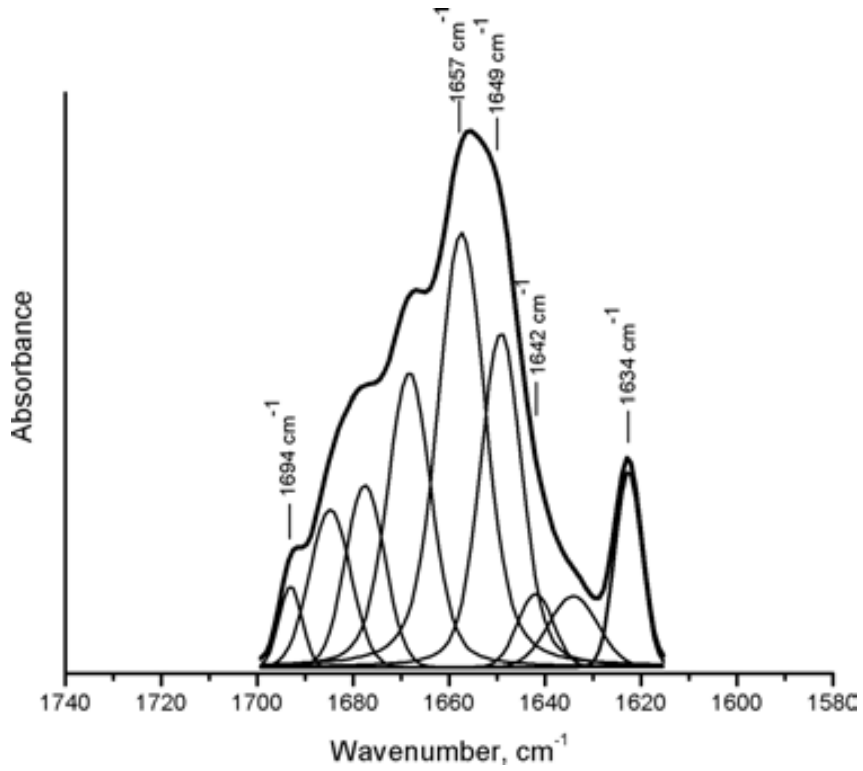


Figure 4.30 Deconvolved spectra and curve fitting of amide I region of hNK1R C-terminus in 20 mM sodium phosphate buffer (pH 7.4). The maxima of deconvolved bands are shown.

Figure 4.30 shows a deconvolved and curve-fitted amide I spectra of the absorbance spectrum of figure 4.29 (*solid trace*). Two major bands at 1657 cm^{-1} and 1649 cm^{-1} are revealed as the major amide I band components. Minor components are also detected at 1634, 1642, 1668, 1678, 1686 and 1694 cm^{-1}. The band at 1657 cm^{-1} can be assigned to amide carbonyl groups involved in *α*-helix formation and constitutes about 29% of total hNK1R C-terminus secondary structure, as quantified by curve fitting (table 4.6).

Table 4.6 Secondary structure assignment for hNK1R C-terminus based on curve-fitted FTIR amide I band.

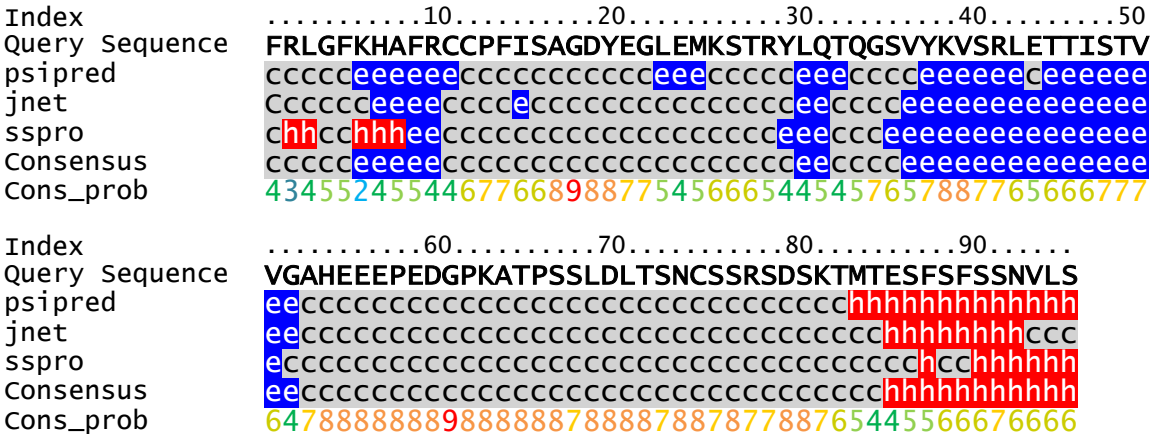
Wavenumber (cm^{-1})	% Area	Assignment
1623	7	*β*-sheet
1634	4	*β* -sheet
1642	3	*β* -sheet
1649	19	random
1657	29	*α*-helix
1668	18	*β* -turn
1678	9	*β* -turn
1686	8	*β* -turn
1694	2	*β* -sheet

The second major component at 1649 cm^{-1} could stand for the presence of unordered structural motifs whose contribution to overall secondary structure is about 19%. Bands which constitute 14% of total structure at 1623 to 1642 cm^{-1} are suggested to correspond to *β*-sheets and four bands that occupy 37% of overall amide I area ranging from 1668 cm^{-1} to 1694 cm^{-1} are assigned to *β*-turns (Byler and Susi, 1986).

4.4.6. Secondary structure prediction for hNK1R C-terminus

The secondary structure of hNK1 C-terminus was predicted using Phyre WEB-based prediction server. It should be pointed out that Phyre not only recognizes the secondary structure of a protein but can also be used as a tool for tertiary structure prediction as far as it is based on protein homology/analogy recognition engine (Kelley and Sternberg, 2009). Figure 4.31 shows secondary structure prediction results generated for hNK1R C-terminus.

A



B

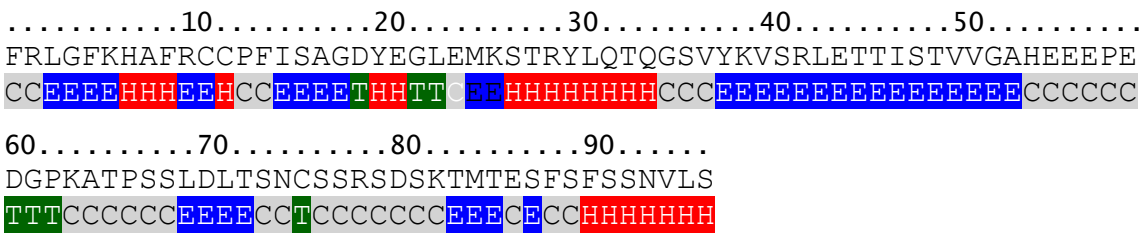


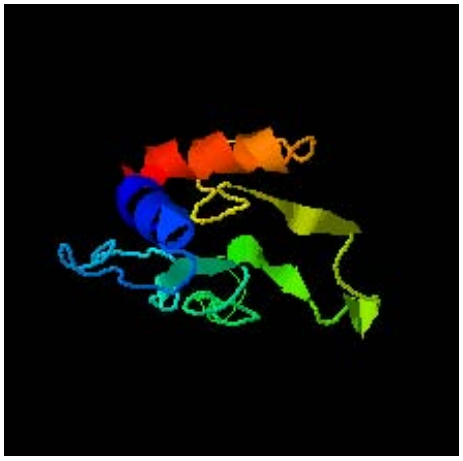
Figure 4.31 *(A)* Secondary structure prediction results for hNK1R C-terminus made by Phyre server and *(B)* SSpro8 prediction server. h − *α*-helix; e − *β*-sheet; T − turns and c − coiled region;. *Number* in cons_prob string is a confidence values at each position of the query for each of the three secondary structure states.

Three independent secondary structure prediction programs were used by Phyre server for the query: Psi-Pred (McGuffin et al., 2000), SSPro (Pollastri et al., 2002) and JNet15 (Cole et al., 2008). Additionally we made a prediction of hNK1R C-terminus secondary structure using extended SSpro8 protocol. It should be pointed out that SSPro and SSpro8 prediction results not always match as they use prediction protocols that are completely different.

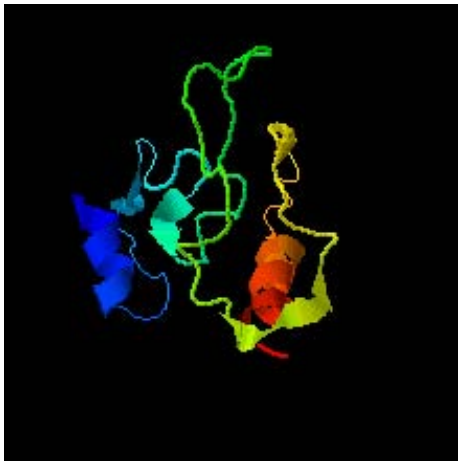
One of the criteria we used to estimate the accuracy of the prediction results was the prediction for the helix VIII spanning from Phe^{312} to Arg^{321}. This small helical hydrophobic fragment was described for many A class GPCRs. As far as helix VIII was detected only by SSpro8 8-class prediction program we gave preference to its results for interpretation of hNK1R C-terminus secondary structure. According to the prediction results the peptide has about 67% of ordered structural elements: about 24% *α*-helix; 36% extended *β*-sheets and 7% *β*-turns. Random coils constitute about 24% of overall secondary structure.

4.4.7. Tertiary structure prediction using I-TASSER

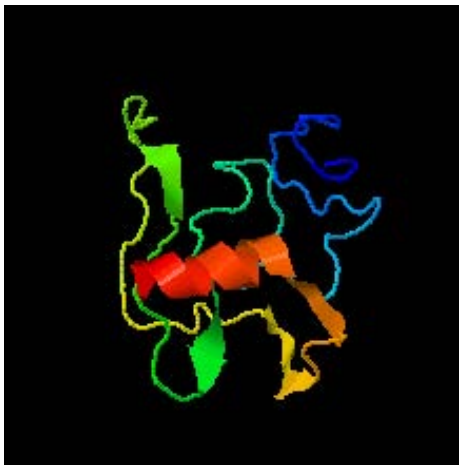
We made a prediction of hNK1R C-terminus tertiary structure using I-TASSER, a hierarchical protein structure modeling method based on the secondary-structure enhanced Profile-Profile threading Alignment (PPA) and the iterative implementation of the Threading ASSEmbly Refinement (TASSER) program. The community-wide Critical Assessment of Structure Prediction (CASP) ranked the I-TASSER as the best protein prediction method in the server section.



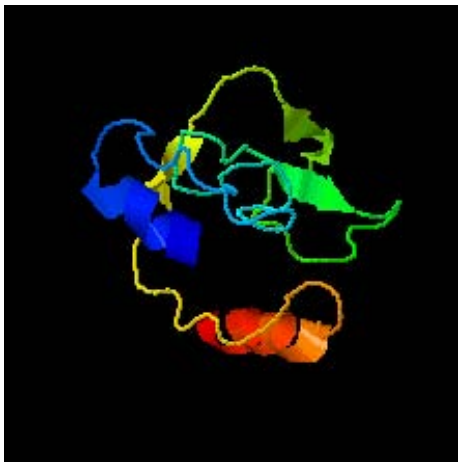
Model 1
C-score=-2.91



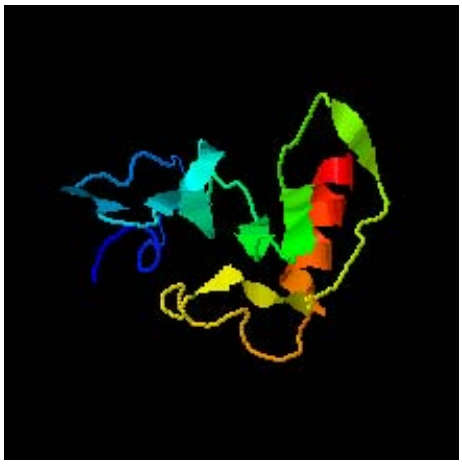
Model 2
C-score=-3.35



Model 3
C-score=-3.69



Model 4
C-score=-3.50



Model 5
C-score=-0.24

Figure 4.32 Tertiary structure prediction results for hNK1R C-terminus made by I-TASSER server. Five generated models are shown. C-score is provided for each model.

Five models were generated by I-TASSER server based on the provided primary sequence of hNK1 C-terminus (figure 4.32). Each of them was analyzed for the best conformity to the experimental results. Model 2 was chosen for further analysis and tertiary structure interpretation as the closest to experiment results obtained from spectroscopic assays. The criteria that were applied for best model selection were the secondary structure elements percentage composition, the prediction of helix VIII and tyrosine allocation able to explain the spectroscopic data.

The results obtained from the prediction of the secondary and tertiary structure for hNK1R C-terminus are in agreement with the results of the spectroscopic studies and will be discussed in the next section.

GENERAL DISCUSSION

5

5.1. Refolding of the hNK1 receptor

Membrane receptors that belong to G protein-coupled receptor family are involved in many physiological processes, including cell signal transduction, sensory signaling, neuronal transmission and hormonal signaling. The tachykinin receptor 1 also known as Neurokinin-1 receptor is a GPCR found in the central and peripheral nervous systems of vertebrates. Substance P is known as a major endogenous ligand for this receptor, although it has some affinity for other tachykinins. NK1R is responsible for many physiological processes, which include pain transmission, exocrine and endocrine secretion, vasodilatation, modulation of cell proliferation, and regulation of the immune and inflammatory responses (Tuluc et al., 2009).

Structural studies as crystallization, NMR and other spectroscopic techniques require big amounts of purified and fully functional receptor. The expression in mammalian cell lines has been a traditional method of choice for production of functional GPCRs, however, it is not widely used for high-level expression due to the high cost of the cultivation media and relatively low protein yields. GPCRs expression in insect cells gives better receptor yields but the high cost and the complexity of this method makes difficult its use for many researchers.

One of the promising strategies to produce recombinant GPCRs for structural studies is an *E. coli* expression system. However, generally, bacterial expression has been hampered by the relatively low yields of GPCRs owing to the toxic effects caused by these 7TM receptors when inserted into bacterial membranes. By contrast, when GPCRs are directed to bacterial inclusion bodies, the highly expressed recombinant receptors are inactive and require refolding (Lundstrom, 2005)

The refolding of membrane receptors and GPCRs in particular is a complicated task that requires screening and adjustment of buffer conditions that can be stabilizing for protein structure upon folding or lead to the formation of misfolded conformations of the protein. The overall efficiency is determined by the two competing processes – refolding and aggregation and a very fine balance should be found in establishing conditions that favor refolding.

In this work we expressed hNK1-366 and hNK1-311 truncated forms of the receptor in *E. coli*. The truncated forms with shortened C-terminus were already used for other GPCRs. As for example, the truncation of C-terminus of hA2aR receptor was necessary to achieve protease resistance and to allow a homogeneous receptor preparation (Weiss and Grisshammer, 2002). The hNK1R-366 upon overexpression in *E. coli* was found in the form of inclusion bodies which may be explained by the toxicity of the receptor for the host cell. Different detergents were tested on their ability to solubilize the extracted IBs. Only strong ionic detergents as SDS and Sarkosyl and high concentration of chaotropic agents (6 M GuHCl, 8 M Urea) were capable of extracting the overexpressed receptor. After purification on His-select nickel affinity column in refolding conditions we obtained significantly less hNK1R-311 than hNK1R-366 which could be indicative of partial hNK1R-311 aggregation in the purification column.

One of the important steps of the obtainment of functional GPCRs from *E. coli* inclusion bodies is the refolding procedure. Considering literature about folding processes, it is common knowledge that physical parameters such as pH, ionic strength, or temperature may have a tremendous effect on the efficacy of *in vitro* folding processes. Furthermore, buffer components may be essential for correct *in vitro* folding. The exact mechanism by which an additive affects the folding of a given protein is difficult to assess. The additives may influence both the solubility and the stability of the unfolded protein, folding intermediates as well as the final native state (Rudolph and Lilie, 1996).

One of the essential components of the refolding system for membrane proteins is a mild detergent capable of keeping the receptor in soluble state. In this work we used CD, UV and fluorescence spectroscopies to monitor the secondary and tertiary hNK1R-366 structure in SDS and DDM and changes produced upon exchange of GuHCl chaotropic agent to mild non-ionic DDM detergent.

The results obtained from CD studies hNK1R-366 shows a decrease of MRE at 222 nm upon exchange of SDS to mild DDM detergent. When refolded in DDM, the receptor presents similar α-helical content as rhodopsin extracted from bovine retinas and solubilized in non-denaturing DDM micelles.

Regarding the results of intrinsic tryptophan fluorescence studies, we observed a blue-shift in the emission spectrum peak of hNK1R-366 at low concentrations of GuHCl. Such changes typical for tryptophan in hydrophobic environment are indicative of protein renaturation. Furthermore, the emission spectra of hNK1R-366 expressed in COS-1 mammalian cells and solubilized in DDM micelles shows very similar emission maximum around 335 nm to that of the receptor refolded from the inclusion bodies. The refolded in physiological buffer hNK1R-366 was prone to aggregate in about 24 hours which was confirmed by changes in turbidity of the samples at 400 nm by UV spectroscopy. The presence of 0.05% DDM was found to stabilize the receptor which was concluded from the absence of Abs at 400 nm.

On-column refolding appears to be a powerful approach for difficult refolding problems and in some cases can be more efficient than other refolding methods. Proteins bound to solid supports, such as IMAC adsorbents, are spatially constrained during the refolding process, preventing them from diffusing toward each other and aggregating when they are in a partially refolded state (Li et al., 2003).

To obtain properly folded receptor, we established a refolding protocol accomplished by several washes applied to denatured protein bound to IMAC purification column. During renaturation process, the strong ionic Sarkosyl detergent used for solubilization of inclusion bodies was stripped by the addition of 5 mM *β*-cyclodextrin. Cyclodextrins, in the presence or absence of detergents, have been reported to suppress aggregate formation during the refolding of a number of proteins (Desai et al., 2006; Cooper, 1992).

Stabilization of the transmembrane helices was ensured by the addition of 0.2% of mild DDM detergent whose ability to stabilize the refolded receptor was clearly seen from the results of spectroscopic assays.

All GPCRs with rare exceptions contain two highly conserved cysteines in the TM3 domain and EC2 loop. It is believed that this pair of Cys residues forms a disulfide bond in most GPCRs. This linkage has been proposed to be important to allow the receptor to attain a normal conformation during synthesis, for normal expression on the cell surface or to maintain normal function, in particular, normal binding and activation (Perlman et al., 1995). For proper disulfide bond formation

during the refolding process a redox system of reduced/oxidized glutathione (GSH/GSSG) is commonly used by many researchers. In our study of refolding of hNK1 receptor, to provide optimal conditions for disulfide bond formation, a redox shuffling system consisting of 1 mM GSH and 0.5 mM GSSG was included in the renaturation buffer.

After refolding on IMAC column both the hNK1R-366 and hNK1R-311 were subjected to radioligand binding assay in order to test their ability to bind [^{3}H]-SP.

We were able to get the saturation binding curve for hNK1R-366 in increasing concentrations of the ligand ranging from 0.05 to 4 nM. The equilibrium dissociation constant K_d for refolded receptor was found to be 1.7 ± 0.6 nM, whereas for full-length hNK1 receptor expressed in COS-1 cells the K_d equals to 0.88 ± 0.26 nM. Using the non-linear curve fitting with an equation based on the law of mass actions with a 1:1 stoichiometry of receptor and ligand we calculated the amount of active hNK1R-366 receptor that was 0.25 ± 0.04 nM.

Unfortunately we were not able to get any saturation binding curve for hNK1R-311 truncated form of the receptor, because it turned out to be incapable of binding any detectable amounts of [^{3}H]-SP. *In vivo* NK1R-311 truncated form is known to be functional and can bind SP, however the binding affinity for the truncated receptor is 10 fold lower than the full length receptor (Fong et al., 1992). On the other hand some authors believe that the amino acid residues between Arg^{311} and Pro^{324} in the C-terminal domain of full length NK1R are not only essential for coupling to G_q but most likely to be involved in the correct folding of the protein, ensuring the optimal conformation of the binding site for SP on the extracellular domains of the receptor (Tuluc et al., 2009). Considering this, the negative radioligand binding assay results could be explained by incorrect folding caused by the lack of 96 residues of the C-terminus of the receptor.

In summary, we have shown the moderate yield expression in *E. coli* and functional refolding in non-denaturing micelles of hNK1R-366 and hNK1R-311 truncated forms of the receptor. The functionality of the refolded hNK1R-366 was established in saturation radioligand-binding assays, where we could characterize the binding of Substance P to the receptor. According to our results, the amount of

the active refolded receptor is about 1% of the total protein the sample. This means that we have a mix of functional and non-functional molecules of the receptor in which the SP binding site is altered. In the future, in order to increase the yield of the functional fraction of the receptor, the refolding conditions should be optimized. As for example, the reconstitution in proteoliposomes could be used to create a favorable environment for stabilizing hNK1 transmembrane domains structure. On the other hand, the binding of SP to hNK1R-366 refolded receptor in nanomolar range is significant and can be considered as a promising result, since until now the intents to produce any detectable amounts of functional NK1 receptor in *E. coli* have been unsuccessful (Bane et al., 2007).

5.2. hNK1 C-terminus structural characterization

C-terminus of GPCRs is a region whose role was underestimated by researchers for a long time and as a result very little information is known about its structure. However, nowadays, GPCRs C-terminus has been recognized as the main domain for the regulation of GPCR functions. Binding motifs potentially involved in protein-protein interactions are specific to each GPCR C-terminus and many splice variants of a given GPCR show sequence variations within the C-terminus domain. It is worth pointing out that, posttranslational modifications as, for example, phosphorylation and palmitoylation occur within the GPCR C-terminus. On the other hand, the proximal region following the seventh transmembrane domain in the GPCR C-terminus or so called helix VIII is known to be crucial for the coupling with G-proteins (Bockaert et al., 2004). Finally, the C-terminus is also particularly important for the control of the equilibrium between the active and inactive forms of GPCRs in the absence of ligands (Claeysen et al., 1999).

Molecular modeling of human Neurokinin-1 receptor tertiary structure based on bovine rhodopsin crystal structure predicted only a small initial antiparallel *β*-sheet, spanning Thr^{339} to Arg^{340} and Gly^{346} to Ser^{347} in C-terminus region. A globular structure was predicted for the rest of the C-terminus residues (Khedkar et al., 2005). The present work is an attempt to give an insight into the

hNK1 C-terminus structure using biochemical methods and different spectroscopic techniques.

Structural and biophysical studies of proteins and peptides generally require high expression levels and therefore a proper expression system should be found in every particular case. The *E. coli* expression system has been successfully used in many studies aiming to obtain large quantities of material for subsequent research (Weiss and Grisshammer, 2002). We expressed hNK1R C-terminus starting from Phe^{312} residue in *E. coli* and purified it using 6X-His tag. Expression results show that the peptide is prone to aggregate in *E. coli* host cytoplasm forming insoluble inclusion bodies which need to be solubilized in strong detergents or chaotropic agents. However, the purified hNK1R C-terminus is soluble in physiological buffers at neutral pH values. At pH values lower than 6 and over 9 insoluble aggregates are formed.

Circular dichroism and FTIR experiments results show that hNK1R C-terminus has clearly defined secondary structure patterns. The quantification based on CD and curve-fitted FTIR spectra are in agreement and reveals from 23 to 29% of helical content for hNK1 C-terminus. Other structural elements as *β*-sheets and *β*-turns constitute about 40% and unordered regions represent from 20 to 35% of hNK1 C-terminus secondary structure.

Fluorescence properties of tyrosines are useful indicators of local microenvironment within a protein. The hNK1 C-terminus has three tyrosine residues, Tyr^{331}, Tyr^{341} and Tyr^{349} located in the proximity of the N-end of the peptide. They are used as intrinsic sensors in UV and fluorescence spectroscopy studies. Absorbance UV spectra recorded for the peptide show abnormal tyrosine red-shifted absorbance band at 292 nm which could be attributed to ionized form of tyrosine. Two other tyrosine absorption bands detected at 277 and 283 nm are characteristic for tyrosine in non-polar environment (Padrós et al., 1982) and can stand for the existence of a structural motif buried inside the globular structure preventing tyrosines from the interaction with polar environment.

Tyrosinate absorption is not likely to appear in ground state at physiological pH values. Therefore, we suggest that such spectral perturbations in tyrosine resulting

in appearance of an anomalous band around 292 nm may be attributed to hydrogen bonded complexes that arise from the proximity of one or more tyrosines to carboxyl groups of glutamic o aspartic residues (Szabo et al., 1979; O'Neil and Hofmann, 1987). The presence of anomalous tyrosine species was also supported by intrinsic fluorescence studies results.

The anomalous tyrosine UV absorption and emission of hNK1 C-terminus can be correlated with the primary amino acid sequence of the peptide rich in aspartic and glutamic residues (figure 5.1), which are known to be good proton acceptors due to conjugate bases of their carboxyl groups. Two regions should be pointed out: the first one starts from Asp^{330} and extends up to Glu^{335} enclosing Tyr^{331} residue; the other one extends from Asp^{366} to Glu^{371}. These two regions rich in proton acceptor residues being in proximity to tyrosyl moieties may cause perturbations in fluorescence emission spectra maxima.

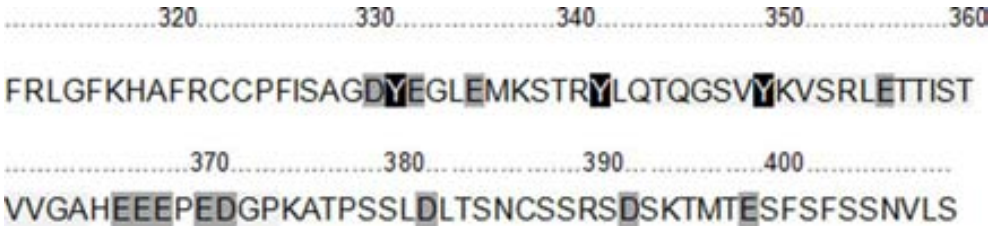


Figure 5.1 Primary sequence of hNK1 C-terminus. Tyrosines (*black*), glutamic and aspartic residues (*dark grey*) are shown.

The analysis of the spectra of GuHCl denatured hNK1R C-terminus made possible the assignment of environment for each tyrosine residue. As far as denaturants affect components of peptides by breaking their secondary and tertiary structures, it may result in the disruption of the interactions between tyrosine phenyl moiety and other residues side chains, hence, normalization of anomalous tyrosine fluorescence should be expected upon unfolding (Lim and Kimura, 1979). However, in case of hNK1R C-terminus the emission component at 345 nm assigned to tyrosinate in polar environment remains even at high GuHCl concentrations.

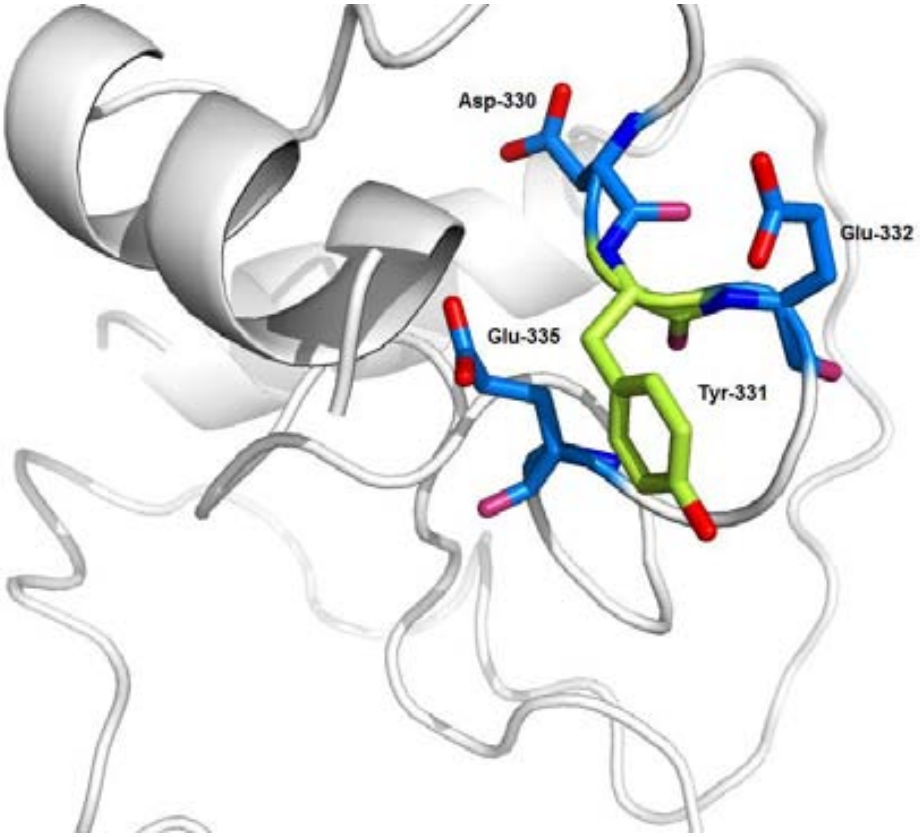


Figure 5.2 Environment of Tyr^{331} residue of hNK1R C-terminus possibly implicated in tyrosinate formation. Oxygens of Glu^{332}, Asp^{330} and Glu^{335} that can become proton acceptor are shown in *light-red*. From a 3D-model of hNK1R C-terminus modeled by I-TASSER server.

As it can be noticed from the primary sequence of hNK1 C-terminus (figure 5.1), Tyr^{331} residue is located between Asp^{330} and Glu^{332} residues and in the immediate vicinity from Glu^{335} residue. The interaction of phenyl moiety hydroxyl with side chain carboxyl of these residues may lead to tyrosinate formation that is not dependent on peptide's tertiary structure changes occurring during folding/unfolding process (figure 5.2). Alternatively, GuHCl may be not capable of breaking the environment of Tyr^{331}. At 6 M GuHCl the emission maxima is at 343 nm and is close to tyrosinate fluorescence in polar environment. The second emission component at 300 nm started to appear on the spectra at as low as 3 M GuHCl concentration. We attributed this band to normal tyrosine fluorescence that possibly originates from Tyr^{341} and Tyr^{349} residues, which become fully exposed to polar environment upon denaturalization (figure 5.3).

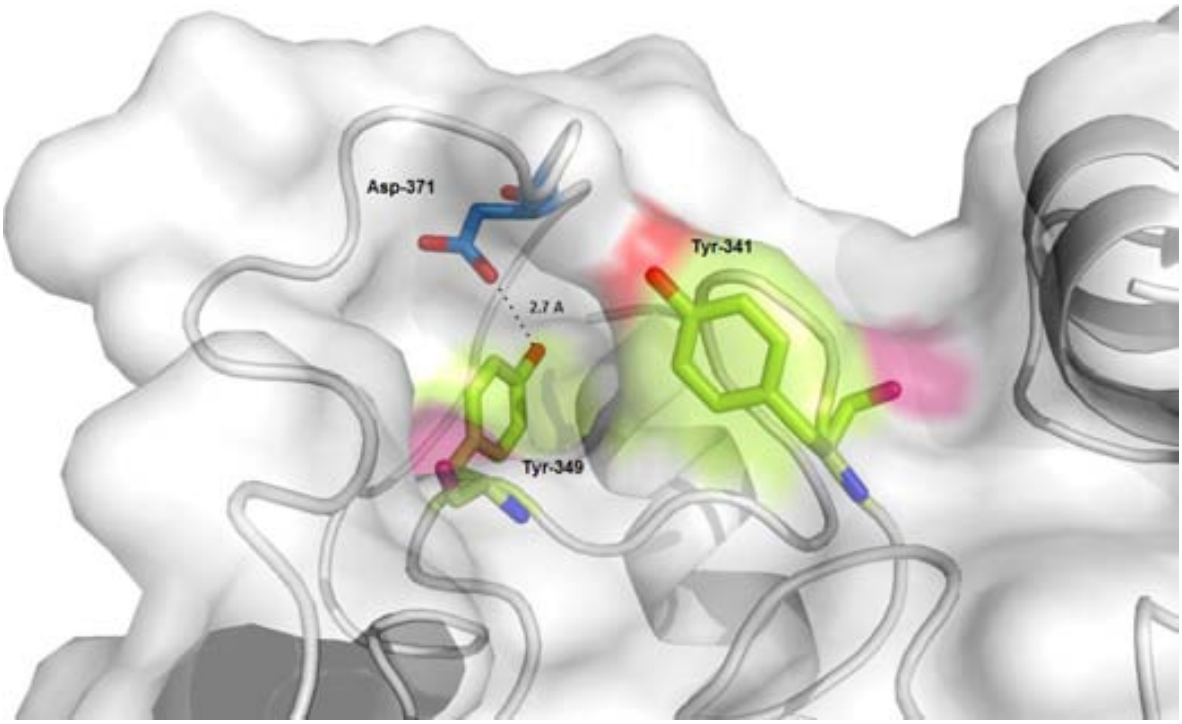


Figure 5.3 Environment of Tyr^{341} and Tyr^{349} residues of hNK1R. Hydrogen bond of 2.7 A is shown between Tyr^{349} and an oxygen of Asp^{371} residue. Tyr^{349} is located in the interior of the peptide, whereas Tyr^{341} is exposed to solvent. From a 3D-model of hNK1R C-terminus modeled by I-TASSER server.

Taking into consideration the results of electrophoretic and spectroscopic studies discussed above we would like to give an insight into hNK1 C-terminus domain organization focusing on its most important structural features. Studies made for C-terminus of other GPCRs showed that their secondary structure may be very different. Fluorescence resonance energy transfer (FRET) studies for β_2-adrenoreceptor C-terminus suggest that it has an extended, relatively unstructured conformation (Granier et al., 2007). On the other hand, C-tail of the human CB1 (cannabinoid 1 receptor) receptor CD spectroscopy assay revealed that the peptide adopts α-helical conformation in negatively charged and zwitterionic detergents, whereas it exhibited the CD signature of unordered structure at low concentration in aqueous solution. 27% helicity was displayed at high peptide concentration suggesting that self-association induces helix formation in the absence of a membrane mimetic (Ahn et al., 2009).

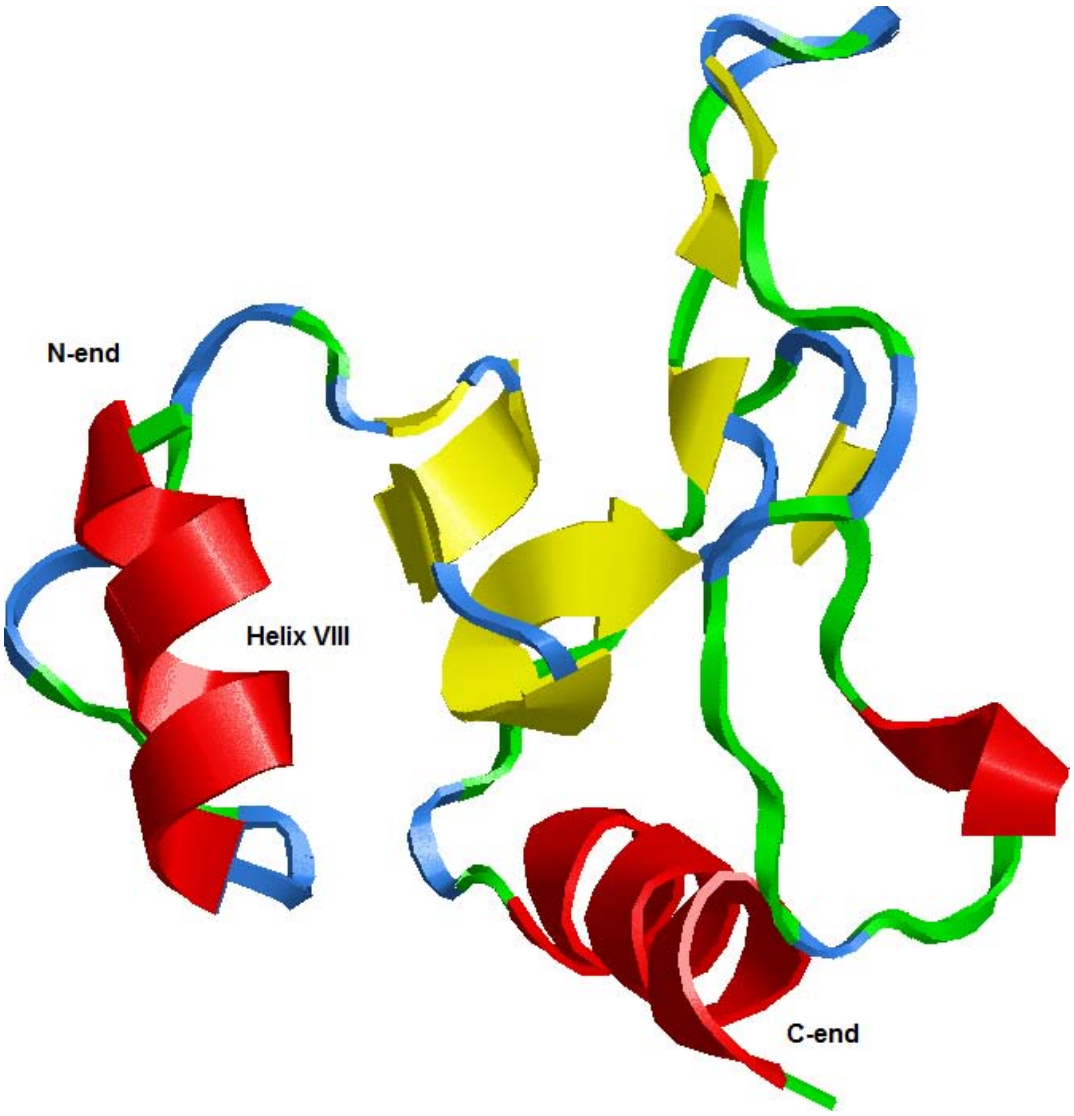


Figure 5.4 3D-model of hNK1R C-terminus modeled by I-TASSER server. Secondary structure elements are colored as follows: *α*-helices are shown in *red; β*-helices are yellow arrows; *β*-turns are *blue* regions and coils are marked with *green* color.

In respect to hNK1R C-terminus secondary structure we made the following assignment: 25% *α*-helix, 27% unordered structure, 48% *β*-sheets and *β*-turns. hNK1R C-terminus has a short helical fragment spanning from Arg^{313} to Arg^{321}, which is about 10% of identified by CD and FTIR studies peptide's *α*-helical content. A similar helical domain, termed helix VIII was described for many A class GPCRs and its hydrophobic residues, structure and relative to TM7 location was determinant for receptor localization leading to robust ligand binding and G protein activation (Ahn et al., 2010). The other 15-17% of *α*-helical content is assigned to

Met^{395} – Ser^{407} region as shown in the I-TASSER generated 3D hNK1R C-terminus model (figure 5.4). According to CD and FTIR results *β*-sheets constitute about 15% of overall secondary structure elements and possibly can be assigned to a region spanning from Val^{348} to Gly^{363} residue. The rest of the secondary structure content may be attributed to coiled and *β*-turns elements. In conclusion, the results discussed above give evidence that hNK1R C-terminus is not an unordered region but has clearly defined secondary and tertiary structures which certainly are tightly related to its multiple functions as receptor desensitization, internalization and recycling.

CONCLUSIONS

6

1. We cloned hNK1R-366, hNK1R-311 truncated form of the receptor and hNK1R C-terminus genes into pET23a vector in order to carry out the expression under T7 promotor in C43 *E. coli* strain.

2. Both truncated hNK1R-366 and hNK1R-311 forms of the receptor, as well as hNK1R C-terminus when expressed in *E. coli*, form non-soluble inclusion bodies in the cytoplasm of the host cell.

3. Detergent screening results shows that hNK1R-366, hNK1R-311 and hNK1R C-terminus inclusion bodies need to be solubilized in strong ionic detergents as SDS and Sarkosyl, or in high concentration of a chaotropic agent as GuHCl. Mild non-ionic detergents as DDM and LAPAO are not capable of solubilizing the IBs.

4. The expression rate of hNK1R-366 is 2.3 times higher than that of the hNK1R-311 when purified in denaturing conditions. After purification on His-select nickel affinity column in refolding conditions we obtained significantly less (3.5 times) hNK1R-311 than hNK1R-366 which could be due to partial hNK1R-311 aggregation in the purification column.

5. CD studies of denatured and refolded hNK1R-366 receptor show that the receptor refolded in mild non-ionic DDM detergent has amount of α-helical similar to that of the rhodopsin extracted from bovine retinas and solubilized in DDM micelles, which could be indicative of correct hNK1R-366 folding.

6. A blue-shift in the emission spectrum peak of hNK1R-366 observed at low concentrations of GuHCl is typical for tryptophan in hydrophobic environment and is indicative of protein renaturation.

7. Emission spectra of hNK1R-366 expressed in the membrane of COS-1 mammalian cells shows very similar emission maximum around 335 nm to that of the receptor refolded from the inclusion bodies. The shift becomes more evident if 0.05% DDM is added to the refolding buffer, indicating that the micellar structure of this non-ionic detergent stabilize the refolded receptor structure.

8. Refolded in physiological buffer hNK1R-366 was prone to aggregate in about 24 hours which was confirmed by changes in turbidity of the samples at 400 nm by UV-vis spectroscopy. The presence of 0.05% DDM was found to stabilize the receptor which was concluded from the absence of absorbance at 400 nm.

9. Refolded hNK1R-366 binds SP with K_d value of 1.7 ± 0.6 nM. The concentration of the active hNKR-366 receptor was 0.25 ± 0.04 nM. The K_d value for refolded receptor appeared to be 2 times higher than for the recombinant receptor expressed in the membrane of COS-1 cells.

10. Refolded truncated hNK1R-311 receptor was unable to bind any detectable amounts of SP. We attribute this to either incorrect receptor folding conditioned by the lack of 96 residues of the C-terminus of the receptor.

11. Circular dichroism and FTIR experiments results show that hNK1R C-terminus has clearly defined secondary structure patterns. The quantification based on CD and curve-fitted FTIR spectra is in agreement and reveals about 25% of helical content for hNK1 C-terminus. Other structural elements as *β*-sheets and *β*-turns constitute about 48% and unordered regions represent 27% of hNK1R C-terminus secondary structure.

12. The UV and fluorescence spectroscopic studies revealed abnormal tyrosine red-shifted absorbance band at 292 nm and intrinsic tyrosine emission at 345 nm which could be attributed to ionized form of tyrosine and possibly arises from the proximity of one or more tyrosines to carboxyl groups of glutamic o aspartic residues.

13. Based on the results of spectroscopic studies as well as on secondary and tertiary structure prediction, we proposed a tridimensional structure model for hNK1R C-terminus.

BIBLIOGRAPHY 7

A

Abrahmsen L., Moks T., Nilsson B., Uhlen M. *Secretion of heterologous gene products to the culture medium of Escherichia coli.* Nucleic Acids Res, 1986. **14**, 7487-7500.

Ahn K. H., Nishiyama A., Mierke D. F., Kendall D. A. *Hydrophobic residues in helix 8 of cannabinoid receptor 1 are critical for structural and functional properties.* Biochemistry, 2010. **49**, 502-511.

Ahn K. H., Pellegrini M., Tsomaia N., Yatawara A., Kendall D. A., Mierke D. F. *Structural analysis of the human cannabinoid receptor one carboxyl-terminus identifies two amphipathic helices.* Biopolymers, 2009. **91**, 565-573.

Attrill H., Harding P. J., Smith E., Ross S., Watts A. *Improved yield of a ligand-binding GPCR expressed in E. coli for structural studies.* Protein Expr Purif, 2009. **64**, 32-38.

B

Bane S. E., Velasquez J. E., Robinson A. S. *Expression and purification of milligram levels of inactive G-protein coupled receptors in E. coli.* Protein Expr Purif, 2007. **52**, 348-355.

Bertin B., Freissmuth M., Breyer R. M., Schutz W., Strosberg A. D., Marullo S. *Functional expression of the human serotonin 5-HT1A receptor in Escherichia coli. Ligand binding properties and interaction with recombinant G protein alpha-subunits.* J Biol Chem, 1992. **267**, 8200-8206.

Bockaert J., Marin P., Dumuis A., Fagni L. *The 'magic tail' of G protein-coupled receptors: an anchorage for functional protein networks.* FEBS Lett, 2003. **546**, 65-72.

Bockaert J. and Pin J. P. *Molecular tinkering of G protein-coupled receptors: an evolutionary success.* EMBO J, 1999. **18**, 1723-1729.

Bockaert J., Roussignol G., Becamel C., Gavarini S., Joubert L., Dumuis A., Fagni L., Marin P. *GPCR-interacting proteins (GIPs): nature and functions.* Biochem Soc Trans, 2004. **32**, 851-855.

Bohm S. K., Khitin L. M., Smeekens S. P., Grady E. F., Payan D. G., Bunnett N. W. *Identification of potential tyrosine-containing endocytic motifs in the carboxyl-tail and seventh transmembrane domain of the neurokinin 1 receptor.* J Biol Chem, 1997. **272**, 2363-2372.

Bonner T. I. *The molecular basis of muscarinic receptor diversity.* Trends Neurosci, 1989. **12,** 148-151

Breyer R. M., Strosberg A. D., Guillet J. G. *Mutational analysis of ligand binding activity of beta 2 adrenergic receptor expressed in Escherichia coli.* EMBO J, 1990. **9**, 2679-2684.

Buck F., Wang W., Harder S., Brathwaite C., Bruhn T. O., Gershengorn M. C. *Juxtamembrane regions in the third intracellular loop of the thyrotropin-releasing hormone receptor type 1 are important for coupling to Gq.* Endocrinology, 2000. **141**, 3717-3722.

Busuttil B. E., Turney K. L., Frauman A. G. *The expression of soluble, full-length, recombinant human TSH receptor in a prokaryotic system.* Protein Expr Purif, 2001. **23**, 369-373.

Byler D. M. and Susi H. *Examination of the secondary structure of proteins by deconvolved FTIR spectra.* Biopolymers, 1986. **25**, 469-487.

C

Cascieri M. A., Ber E., Fong T. M., Sadowski S., Bansal A., Swain C., Seward E., Frances B., Burns D., Strader C. D. *Characterization of the binding of a potent, selective, radioiodinated antagonist to the human neurokinin-1 receptor.* Mol Pharmacol, 1992. **42**, 458-463.

Chapot M. P., Eshdat Y., Marullo S., Guillet J. G., Charbit A., Strosberg A. D., Delavier-Klutchko C. *Localization and characterization of three different beta-adrenergic receptors expressed in Escherichia coli.* Eur J Biochem, 1990. **187**, 137-144.

Chin C. N., von Heijne G., de Gier J. W. *Membrane proteins: shaping up.* Trends Biochem Sci, 2002. **27**, 4-231.

Claeysen S., Sebben M., Becamel C., Bockaert J., Dumuis A. *Novel Brain-Specific 5-HT4 Receptor Splice Variants Show Marked Constitutive Activity: Role of the C-Terminal Intracellular Domain.* Mol Pharmacol, 1999. **55**, 910-920.

Coge F. and Regoli D. *NK1 and NK2 receptors are similar in man and rabbit.* Neuropeptides, 1994. **26,** 385-90.

Cole C., Barber J. D., Barton G. J. *The Jpred 3 secondary structure prediction server.* Nucleic Acids Res, 2008. 36 (Web server issue), W197–W201.

Cooper A. *Effect of cyclodextrins on the thermal stability of globular proteins.* Am Chem Soc, 1992. **114**, 9208-9214

D

Datar P., Srivastava S., Coutinho E., Govil G. *Substance P: structure, function, and therapeutics.* Curr Top Med Chem, 2004. **4**, 75-103.

Dave N., Lórenz-Fonfría V. A., Villaverde J., Lemonnier R., Leblanc G., Padrós E. *Study of amide-proton exchange of Escherichia coli melibiose permease by attenuated total reflection-Fourier transform infrared spectroscopy: evidence of structure modulation by substrate binding.* J Biol Chem, 2002. **277**, 3380-3387.

Desai A., Lee C., Sharma L., Sharma A. *Lysozyme refolding with cyclodextrins: structure-activity relationship.* Biochimie, 2006. **88**,1435-1445.

Dixon R. A. F., Strader C. D., Sigal I. S., Allen R. C. *Structure and function of G-protein coupled receptors.* Annu Rep Med Chem ed, 1988. **23,** 221-233.

Douglas S. D. and Leeman S. E. *Neurokinin-1 receptor: functional significance in the immune system in reference to selected infections and inflammation.* Ann N Y Acad Sci, 2010. **1217**, 83-95.

E

Ebner K. and Singewald N. *The role of substance P in stress and anxiety responses.* Amino Acids, 2006. **31**, 251-272.

F

Fahnert B., Lilie H., Neubauer P. *Inclusion bodies: formation and utilisation.* Adv Biochem Eng Biotechnol, 2004. **89**, 93-142.

Fong T. M., Anderson S. A., Yu H., Huang R. R., Strader C. D. *Differential activation of intracellular effector by two isoforms of human neurokinin-1 receptor.* Mol Pharmacol, 1992. **41**, 24-30.

Fong T. M., Yu H., Huang R. R., Strader C. D. *The extracellular domain of the neurokinin-1 receptor is required for high-affinity binding of peptides.* Biochemistry, 1992. **31**, 11806-11811.

Fotiadis D., Liang Y., Filipek S., Saperstein D. A., Engel A., Palczewski K. *Atomic-force microscopy: Rhodopsin dimers in native disc membranes.* Nature, 2003. **421**, 127-128.

Foroutan A., Lazarova T., Padrós E. *Study of Membrane-Induced Conformations of Substance P: Detection of Extended Polyproline II Helix Conformation.* J of Phys Chem, 2011. **115**, 3622-3631

Furukawa H. and Haga T. *Expression of functional M2 muscarinic acetylcholine receptor in Escherichia coli.* J Biochem, 2000. **127**, 151-161.

G

Gao Z., Ni Y., Szabo G., Linden J. *Palmitoylation of the recombinant human A1 adenosine receptor: enhanced proteolysis of palmitoylation-deficient mutant receptors.* Biochem J, 1999. **342**, 387-395.

Gibbins I. L., Furness J. B., Costa M., MacIntyre I., Hillyard C. J., Girgis S. *Co-localization of calcitonin gene-related peptide-like immunoreactivity with substance P in cutaneous, vascular and visceral sensory neurons of guinea pigs.* Neurosci Lett, 1985. **57**, 125-130.

Gimpl G., Klein U., Reilander H., Fahrenholz F. *Expression of the human oxytocin receptor in baculovirus-infected insect cells: high-affinity binding is induced by a cholesterol-cyclodextrin complex.* Biochemistry, 1995. **34**, 13794-13801.

Granier S., Kim S., Shafer A., Ratnala V., Fung J., Zare R., Kobilka B. *Structure and conformational changes in the C-terminal domain of the beta2-adrenoceptor: insights from fluorescence resonance energy transfer studies.* Biol Chem, 2007. **282**,13895-13905.

Graziani M., Finazzi-Agro A., Rotillo G., Barra D., Mondovi B. *Parsley plastocyanin. Possible presence of sulfhydryl and tyrosine in the copper environment.* Biochemistry, 1974. **13**, 804–809.

Greenfield N. *Using circular dichroism spectra to estimate protein secondary structure.* Nat Protoc, 2006. **1**, 2876-2890.

H

Haendler B., Hechler U., Becker A., Schleuning W. D. *Expression of human endothelin receptor ETB by Escherichia coli transformants.* Biochem Biophys Res Commun, 1993. **191**, 633-638.

Hampe W., Voss R. H., Haase W., Boege F., Michel H., Reilander H. *Engineering of a proteolytically stable human beta 2-adrenergic receptor/maltose-binding protein fusion and production of the chimeric protein in Escherichia coli and baculovirus-infected insect cells.* J Biotechnol, 2000. **77**, 219-234.

Hasegawa J., Loh H. H., Lee N. M. *Lipid requirement for mu opioid receptor binding.* J Neurochem, 1987. **49**, 1007-1012.

Hesketh P. J. *Potential role of the NK1 receptor antagonists in chemotherapy-induced nausea and vomiting.* Support Care Cancer, 2001. **9**, 350-354.

Hill R. A. and Sillence M. N. *Improved membrane isolation in the purification of beta 2-adrenoceptors from transgenic Escherichia coli.* Protein Expr Purif, 1997. **10**, 162-167.

Hoffmann M., Verzijl D., Lundstrom K., Simmen U., Alewijnse A. E., Timmerman H., Leurs R. *Recombinant Semliki Forest virus for over-expression and pharmacological characterisation of the histamine H(2) receptor in mammalian cells.* Eur J Pharmacol, 2001. **427**, 105-114.

Hokfelt T., Holets V. R., Staines W., Meister B., Melander T., Schalling M., Schultzberg M., Freedman J., Bjorklund H., Olson L. and et al. *Coexistence of neuronal messengers-an overview.* Prog Brain Res, 1986. **68**, 33-70.

Holzer P. *Capsaicin: cellular targets, mechanisms of action, and selectivity for thin sensory neurons.* Pharmacol Rev, 1991. **43**, 143-201.

Huang R. R., Yu H., Strader C. D., Fong T. M. *Interaction of substance P with the second and seventh transmembrane domains of the neurokinin-1 receptor.* Biochemistry, 1994. **33**, 3007-3013.

Hunter J. C., Hannah P. A., Maggio J. E. *The regional distribution of kassinin-like immunoreactivity in central and peripheral tissues of the cat.* Brain Res, 1985. **341**, 228-232.

Huston J. P., Hasenohrl R. U., Boix F., Gerhardt P., Schwarting R. K. *Sequence-specific effects of neurokinin substance P on memory, reinforcement, and brain dopamine activity.* Psychopharmacology (Berl), 1993. **112**, 147-162.

I

Ingi T., Kitajima Y., Minamitake Y., Nakanishi S. *Characterization of ligand-binding properties and selectivities of three rat tachykinin receptors by transfection and functional expression of their cloned cDNAs in mammalian cells.* J Pharmacol Exp Ther, 1991. **259**, 968-975.

J

Jockers R., Linder M. E., Hohenegger M., Nanoff C., Bertin B., Strosberg A. D., Marullo S., Freissmuth M. *Species difference in the G protein selectivity of the human and bovine A1-adenosine receptor.* J Biol Chem, 1994. **269**, 32077-32084.

Jordano J., Barbero J. L., Montero F., Franco L. *Fluorescence of histones H1. A tyrosinate-like fluorescence emission in Ceratitis capitata H1 at neutral pH values.* J Biol Chem, 1983. **258**, 315-320.

K

Kelley L. A. and Sternberg M. J. E. *Protein structure prediction on the Web: a case study using the Phyre server.* Nature Protocols, 2009. **4**, 363-371

Khedkar S. A., Malde A. K., Coutinho E. C. *Modeling Human Neurokinin-1 Receptor Structure Using the Crystal Structure of Bovine Rhodopsin.* Internet Electron J Mol, 2005. **4**, 329-341.

Kiefer H. *In vitro folding of alpha-helical membrane proteins.* Biochim Biophys Acta, 2003. **1610**, 57-62.

Kiefer H., Krieger J., Olszewski J. D., Von Heijne G., Prestwich G. D., Breer H. *Expression of an olfactory receptor in Escherichia coli: purification, reconstitution, and ligand binding.* Biochemistry, 1996. **35**, 16077-16084.

Kiefhaber T., Rudolph R., Kohler H. H., Buchner J. *Protein aggregation in vitro and in vivo: a quantitative model of the kinetic competition between folding and aggregation.* Biotechnology (NY), 1991. **9**, 825-829.

Kobilka B. and Schertler G. F. *New G-protein-coupled receptor crystal structures: insights and limitations.* Trends Pharmacol Sci, 2008. **29**, 79-83.

Kobilka B. K. *G protein coupled receptor structure and activation.* Biochim Biophys Acta, 2007. **1768**, 794-807.

Kotani H., Hoshimaru M., Nawa H., Nakanishi S. *Structure and gene organization of bovine neuromedin K precursor.* Proc Natl Acad Sci U S A, 1986. **83**, 7074-7078.

Krause J. E., Takeda Y., Hershey A. D. *Structure, functions, and mechanisms of substance P receptor action.* J Invest Dermatol, 1992. **98**, 2S-7S.

Kris R. M., South V., Saltzman A., Felder S., Ricca G. A., Jaye M., Huebner K., Kagan J., Croce C. M., Schlessinger J. *Cloning and expression of the human substance K receptor and analysis of its role in mitogenesis.* Cell Growth Differ, 1991. **2**, 15-22.

L

Lacatena R. M., Cellini A., Scavizzi F., Tocchini-Valentini G. P. *Topological analysis of the human beta 2-adrenergic receptor expressed in Escherichia coli.* Proc Natl Acad Sci U S A, 1994. **91**, 10521-10525.

Lai J. P., Ho W. Z., Kilpatrick L. E., Wang X., Tuluc F., Korchak H. M., Douglas S. D. *Full-length and truncated neurokinin-1 receptor expression and function during monocyte/macrophage differentiation.* Proc Natl Acad Sci (USA), 2006. **103**, 7771-7776.

Li M., Su Z. G., Janson J. C. *In vitro protein refolding by chromatographic procedures.* Protein Expression and Purification, 2004. **33**, 1–10

Liang Y., Fotiadis D., Filipek S., Saperstein D. A., Palczewski K., Engel A. *Organization of the G protein-coupled receptors rhodopsin and opsin in native membranes.* J Biol Chem, 2003. **278**, 21655-21662.

Lilie H., Schwarz E., Rudolph R. *Advances in refolding of proteins produced in E. coli.* Curr Opin Biotechnol, 1998. **9**, 497-501.

Lim B. and Kimura T. *Conformation-associated anomalous tyrosine fluorescence of adrenodoxin.* J Biol Chem, 1979. **255**, 2440-2444.

Lohse M. J. *Stable overexpression of human beta 2-adrenergic receptors in mammalian cells.* Naunyn Schmiedebergs Arch Pharmacol, 1992. **345**, 444-451.

London E. and Khorana H. G. *Denaturation and renaturation of bacteriorhodopsin in detergents and lipid-detergent mixtures.* J Biol Chem, 1982. **257**, 7003-7011.

Lundstrom K. *Structural genomics of GPCRs.*TRENDS in Biotechnology, 2005. **23**, 103-108

Lundstrom K., Mills A., Buell G., Allet E., Adami N., Liljestrom P. *High-level expression of the human neurokinin-1 receptor in mammalian cell lines using the Semliki Forest virus expression system.* Eur J Biochem, 1994. **224**, 917-921.

Lundstrom K., Wagner R., Reinhart C., Desmyter A., Cherouati N., Magnin T., Zeder-Lutz G., Courtot M., Prual C., Andre N., Hassaine G., Michel H., Cambillau C., Pattus F. *Structural genomics on membrane proteins: comparison of more than 100 GPCRs in 3 expression systems.* J Struct Funct Genomics, 2006. **7**, 77-91.

M

Maggi C. A. T*achykinin receptors and receptor subtypes.* Gen Pharmacol, 1995. **26**, 911-44.

Malherbe P., Kratzeisen C., Lundstrom K., Richards J. G., Faull R. L., Mutel V. *Cloning and functional expression of alternative spliced variants of the human metabotropic glutamate receptor 8.* Brain Res Mol Brain Res, 1999. **67**, 201-210.

Manzini S., Conti S., Maggi C. A., Abelli L., Somma V., Del Bianco E., Geppetti P. *Regional differences in the motor and inflammatory responses to capsaicin in guinea pig airways. Correlation with content and release of substance P-like immunoreactivity.* Am Rev Respir Dis, 1989. **140**, 936-941.

Marullo S., Delavier-Klutchko C., Eshdat Y., Strosberg A. D., Emorine L. *Human beta 2-adrenergic receptors expressed in Escherichia coli membranes retain their pharmacological properties.* Proc Natl Acad Sci U S A, 1988. **85**, 7551-7555.

Masu Y., Nakayama K., Tamaki H., Harada Y., Kuno M., Nakanishi S. cDNA cloning of bovine substance-K receptor through oocyte expression system. Nature, 1987. **329,** 836-938.

McAllister G., Charlesworth A., Snodin C., Beer M. S., Noble A. J., Middlemiss D. N., Iversen L. L., Whiting P. *Molecular cloning of a serotonin receptor from human brain (5HT1E): a fifth 5HT1-like subtype.* Proc Natl Acad Sci U S A, 1992. **89**, 5517-5521.

McGuffin L. J., Bryson K., Jones D. T. *The PSIPRED protein structure prediction server.* Bioinformatics, 2000. **16**, 404–405.

Michalke K., Huyghe C., Lichiere J., Graviere M. E., Siponen M., Sciara G., Lepaul I., Wagner R., Magg C., Rudolph R., Cambillau C., Desmyter A. *Mammalian G protein-coupled receptor expression in Escherichia coli: II. Refolding and biophysical characterization of mouse cannabinoid receptor 1 and human parathyroid hormone receptor 1.* Anal Biochem, 2010. **401**, 74-80.

Middelberg A. *Preparative protein refolding.* TRENDS in Biotechnology, 2002. **20**, 437-443.

Migeon J. C. and Nathanson N. M. *Differential regulation of cAMP-mediated gene transcription by m1 and m4 muscarinic acetylcholine receptors. Preferential coupling of m4 receptors to Gi alpha-2.* J Biol Chem, 1994. **269**, 9767-9773.

Mobarec J. C. and Filizola M. *Advances in the Development and Application of Computational Methodologies for Structural Modeling of G-Protein Coupled Receptors.* Expert Opin Drug Discov, 2008. **3**, 343-355.

Molina D., Cornvik T., Eshaghi S., Haeggström J. Z. Nordlund P., Ignatushchenko Sabet M*. Engineering membrane protein overproduction in Escherichia coli.* Protein Sci, 2008. **17**, 673–680.

Montserret R., McLeish M. J., Böckmann A., Geourjon C., Penin F. *Involvement of electrostatic interactions in the mechanism of peptide folding induced by sodium dodecyl sulfate binding*. Biochemistry, 2000. **39**, 8362-8373.

Munch G., Walker P., Shine J., Herzog H. *Ligand binding analysis of human neuropeptide Y1 receptor mutants expressed in E. coli.* Receptors Channels, 1995. **3**, 291-297.

Muñoz M. and Covenas R. *Neurokinin-1 receptor: a new promising target in the treatment of cancer.* Discov Med, 2010. **10**, 305-313.

Muñoz M., Gonzalez-Ortega A., Covenas R. *The NK-1 receptor is expressed in human leukemia and is involved in the antitumor action of aprepitant and other NK-1 receptor antagonists on acute lymphoblastic leukemia cell lines.* Invest New Drugs, 2012. **30**, 529-540.

N

Nakanishi A. *Mammalian tachykinin receptors.* Annu. Rev. Neurosci., 1991. **14**, 123-36.

Nakanishi S., Ohkubo H., Kakizuka A., Yokota Y., Shigemoto R., Sasai Y., Takumi T. *Molecular characterization of mammalian tachykinin receptors and a possible epithelial potassium channel.* Recent Prog. Hormone Res, 1990. **46**, 59-84.

Nawa H., Hirose T., Takashima H., Inayama S., Nakanishi S. *Nucleotide sequences of cloned cDNAs for two types of bovine brain substance P precursor.* Nature, 1983. **306**, 32-36.

Needham M., Egerton M., Millest A., Evans S., Popplewell M., Cerillo G., McPheat J., Monk A., Jack A., Johnstone D. and et al. *Further development of the locus control region/murine erythroleukemia expression system: high level expression and characterization of recombinant human calcitonin receptor.* Protein Expr Purif, 1995. **6**, 124-131.

Nunez M. T. and Glass J. *Reconstitution of the transferrin receptor in lipid vesicles. Effect of cholesterol on the binding of transferrin.* Biochemistry, 1982. **21**, 4139-4143.

O

O'Neil J. D. and Hofmann T. *Tyrosine and tyrosinate fluorescence of pig intestinal Ca^{2+}-binding protein.* Biochem J, 1987. **243**, 611–615.

Oprian D. D., Molday R. S., Kaufman R. J., Khorana H. G. *Expression of a synthetic bovine rhodopsin gene in monkey kidney cells.* Proc Natl Acad Sci U S A, 1987. **84**, 8874-8878.

P

Pace C. N., Shirley B. A., McNutt M. and Gajiwala K. *Forces contributing to the conformational stability of proteins.* FASEB J, 1996. **10**, 75-83.

Padrós E., Morros A., Mañosa J., Duñach M. *The state of tyrosine and phenylalanine residues in proteins analyzed by fourth-derivative spectrophotometry. Histone H1 and ribonuclease A*. Eur J Biochem, 1982. **127,** 117-122.

Park S. W., Yan Y. P., Satriotomo I., Vemuganti R., Dempsey R. J. *Substance P is a promoter of adult neural progenitor cell proliferation under normal and ischemic conditions.* J Neurosurg, 2007. **107**, 593-599.

Patra A. K., Mukhopadhyay R., Mukhija R., Krishnan A., Garg L. C., Panda A. K. *Optimization of inclusion body solubilization and renaturation of recombinant human growth hormone from Escherichia coli.* Protein Expr Purif, 2000. **18**, 182-192.

Peeters M. C., van Westen G. J., Li Q., IJzerman A. P. *Importance of the extracellular loops in G protein-coupled receptors for ligand recognition and receptor activation.* Trends Pharmacol Sci, 2010.

Peralta E. G., Ashkenazi A., Winslow J. W., Ramachandran J., Capon D. J. *Differential regulation of PI hydrolysis and adenylyl cyclase by muscarinic receptor subtypes.* Nature, 1988. **334**, 434-437.

Perlman J. H, Wang W., Nussenzveig D. R., Gershengorn M. C. *A Disulfide Bond between Conserved Extracellular Cysteines in the Thyrotropin-releasing Hormone Receptor Is Critical for Binding.* J of Biol Chem, 1995. **270**, 24682-24685.

Pollastri G., Przybylski D., Rost B., Baldi P. *Improving the prediction of protein secondary structure in three and eight classes using recurrent neural networks and profiles.* Proteins, 2002. **47**, 228–235.

Popot J. L., Engelman D. M. *Helical membrane protein folding, stability, and evolution.* Annu Rev Biochem, 2000. **69**, 881-922.

Pradier L., Ménager J., Le Guern J., Bock M. D., Heuillet E., Fardin V., Garret C., Doble A., Mayaux J. F. *Septide: an agonist for the NK1 receptor acting at a site distinct from substance P.* Mol Pharmacol, 1994. **45**, 287-93.

Prendergast F. G., Hampton P. D., Jones B. *Characteristics of tyrosinate fluorescence emission in alpha- and beta-purothionins.* Biochemistry, 1984. **23**, 6690-6697.

R

Raddatz R., Crankshaw C. L., Snider R. M., Krause J. E. *Similar rates of phosphatidylinositol hydrolysis following activation of wild-type and truncated rat neurokinin-1 receptors*. J Neurochem, 1995. **64**, 1183-1191.

Reeves P. J., Kim J. M., Khorana H. G. *Structure and function in rhodopsin: a tetracycline-inducible system in stable mammalian cell lines for high-level expression of opsin mutants.* Proc Natl Acad Sci U S A, 2002. **99**, 13413-13418.

Regoli D., and Fauchére J. L. *Receptors and antagonists for substance P and related peptides.* Pharmacol Rev., 1994. **46**, 99-551.

Regoli D., Escher E., Mizrahi J. *Substance P - structure-activity studies and the development of antagonists.* Pharmacology, 1984. **28**, 301-320.

Ren H., Yu D., Ge B., Cook B., Xu Z., Zhang S. *High-level production, solubilization and purification of synthetic human GPCR chemokine receptors CCR5, CCR3, CXCR4 and CX3CR1.* PLoS One, 2009. **4**, e4509

Rogl H., Kosemund K., Kuhlbrandt W., Collinson I. *Refolding of Escherichia coli produced membrane protein inclusion bodies immobilised by nickel chelating chromatography.* FEBS Lett, 1998. **432**, 21-26.

Rosenbaum D. M., Rasmussen S. G., Kobilka B. K. *The structure and function of G-protein-coupled receptors.* Nature, 2009. **459**, 356-363.

Roy A., Kucukural A., Zhang Y. *I-TASSER: a unified platform for automated protein structure and function prediction.* Nature Protocols, 2010. **5**, 725-738.

Rudolph R. and Lilie H. *In vitro folding of inclusion body proteins.* FASEB J, 1996. **10**, 49-56.

S

Sarkar C. A., Dodevski I., Kenig M., Dudli S., Mohr A., Hermans E., Pluckthun A. *Directed evolution of a G protein-coupled receptor for expression, stability, and binding selectivity.* Proc Natl Acad Sci U S A, 2008. **105**, 14808-14813.

Sarramegna V., Talmont F., Demange P., Milon A. *Heterologous expression of G-protein-coupled receptors: comparison of expression systems from the standpoint of large-scale production and purification.* Cell Mol Life Sci, 2003. **60**, 1529-1546.

Satake H. and Kawada T. *Overview of the primary structure, tissue-distribution, and functions of tachykinins and their receptors.* Curr Drug Targets, 2006. **7**, 963-974.

Saudou F., Boschert U., Amlaiky N., Plassat J. L., Hen R. *A family of Drosophila serotonin receptors with distinct intracellular signalling properties and expression patterns.* EMBO J, 1992. **11**, 7-17.

Schmidlin F., Roosterman D., Bunnett N. *The third intracellular loop and carboxyl tail of neurokinin 1 and 3 receptors determine interactions with beta-arrestins.* Am J Physiol Cell Physiol, 2003. **285,** 945-958.

Schmidt P., Lindner D., Montag C., Berndt S., Beck-Sickinger A. G., Rudolph R., Huster, D. *Prokaryotic expression, in vitro folding, and molecular pharmacological characterization of the neuropeptide Y receptor type 2.* Biotechnol Prog, 2009. **25**, 1732-1739.

Schroder-Tittmann K., Bosse-Doenecke E., Reedtz-Runge S., Ihling C., Sinz A., Tittmann K., Rudolph R. *Recombinant expression, in vitro refolding, and biophysical characterization of the human glucagon-like peptide-1 receptor.* Biochemistry, 2010. **49**, 7956-7965.

Schulein R., Rutz C., Rosenthal W. *Membrane targeting and determination of transmembrane topology of the human vasopressin V2 receptor.* J Biol Chem, 1996. **271**, 28844-28852.

Schwartz T. W. *Locating ligand-binding sites in 7TM receptors by protein engineering.* Curr Opin Biotechnol, 1994. **5**, 434-444.

Scrofani S. D., Fabri L. J., Xu P., Maccarone P., Nash A. D. *Purification and refolding of vascular endothelial growth factor-B.* Protein Sci, 2000. **9**, 2018-2025.

Selkoe D. J. *Folding proteins in fatal ways.* Nature, 2003. **426**, 900-904.

Sharkey K. A., Williams R. G., Schultzberg M., Dockray G. J. *Sensory substance P-innervation of the urinary bladder: possible site of action of capsaicin in causing urine retention in rats.* Neuroscience, 1983. **10**, 861-868.

Song H. S., Lee S. H., Oh E. H., Park T. H. *Expression, solubilization and purification of a human olfactory receptor from Escherichia coli.* Curr Microbiol, 2009. **59**, 309-314.

Sreerama N. and Woody R. W. *A Self-Consistent Method for the Analysis of Protein Secondary Structure from Circular Dichroism.* Anal Biochem, 1993. **209**, 32-44.

Stanasila L., Lim W. K., Neubig R. R., Pattus F. *Coupling efficacy and selectivity of the human mu-opioid receptor expressed as receptor-Galpha fusion proteins in Escherichia coli.* J Neurochem, 2000. **75**, 1190-1199.

Stanasila L., Perez J. B., Vogel H., Cotecchia S. *Oligomerization of the alpha 1a- and alpha 1b-adrenergic receptor subtypes. Potential implications in receptor internalization.* J Biol Chem, 2003. **278**, 40239-40251.

Steiner D. F., Smeekens S. P., Ohagi S., Chan S. J. *The new enzymology of precursor processing endoproteases.* J Biol Chem, 1992. **267**, 23435-23438.

Surrey T. and Jähnig F. *Refolding and oriented insertion of a membrane protein into a lipid bilayer.* Proc Natl Acad Sci, 1992. **89**, 7457–7461.

Szabo A. G., Lynn K., Krajcarski D., Rayner D. M. *Tyrosine fluorescence at 345 nm in proteins lacking tryptophan.* Journal of Luminescence, 1979. **18**, 582-586.

T

Takahashi K., Tanaka A., Hara M., Nakanishi S. *The primary structure and gene organization of human substance P and neuromedin K receptors.* Eur J Biochem, 1992. **204**, 1025-1033.

Takeda K. and Moriyama Y. J. *Circular dichroism studies on helical structure preferences of amino acid residues of proteins caused by sodium dodecyl sulfate.* Protein Chem, 1990. **9**, 573-582.

Takeda Y., Chou K. B., Takeda J., Sachais B. S., Krause J. E. *Molecular cloning, structural characterization and functional expression of the human substance P receptor.* Biochem Biophys Res Commun, 1991. **179**, 1232-1240.

Tate C. G. and Grisshammer R. *Heterologous expression of G-protein-coupled receptors.* Trends Biotechnol, 1996. **14**, 426-430.

Tebben A. J. and Schnur D. M. *Beyond rhodopsin: G protein-coupled receptor structure and modeling incorporating the beta2-adrenergic and adenosine A(2A) crystal structures.* Methods Mol Biol, 2011. **672**, 359-386.

Tousignant C., Guillemette G., Regoli D. *Binding sites for [3H][Sar9, Met(O2)11]substance P in rat brain and guinea pig ileum.* Brain Res, 1991. **560**, 1-11.

Townsend-Nicholson A. and Shine J. *Molecular cloning and characterisation of a human brain A1 adenosine receptor cDNA.* Brain Res Mol Brain Res, 1992. **16**, 365-370.

Tuluc F., Lai J. P., Kilpatrick L. E., Evans D. L., Douglas S. D. *Neurokinin 1 receptor isoforms and the control of innate immunity.* Trends Immunol, 2009. **30**, 271-276.

Turcatti G., Ceszkowski K., Chollet A. *Biochemical characterization and solubilization of human NK2 receptor expressed in Chinese hamster ovary cells.* J Recept Res, 1993. **13**, 639-652

W

Walker P., Munoz M., Combe M. C., Grouzmann E., Herzog H., Selbie L., Shine J., Brunner H. R., Waeber B., Wittek R. *High level expression of human neuropeptide Y receptors in mammalian cells infected with a recombinant vaccinia virus.* Mol Cell Endocrinol, 1993. **91**, 107-112.

Weiss H. M. and Grisshammer R. *Purification and characterization of the human adenosine A(2a) receptor functionally expressed in Escherichia coli.* Eur J Biochem, 2002. **269**, 82-92.

White J. F., Trinh L. B., Shiloach J., Grisshammer R. *Automated large-scale purification of a G protein-coupled receptor for neurotensin*. FEBS Lett, 2004. **564**, 289–293.

Y

Yasuda M., Murakami Y., Sowa A., Ogino H., Ishikawa H. *Effect of additives on refolding of a denatured protein.* Biotechnol Prog, 1998. **14**, 601-606.

Z

Zahn R., von Schroetter C., Wuthrich K. *Human prion proteins expressed in Escherichia coli and purified by high-affinity column refolding.* FEBS Lett, 1997. **417**, 400-404.

Zeng F. Y. and Wess J. *Identification and molecular characterization of m3 muscarinic receptor dimers.* J Biol Chem, 1999. **274**, 19487-19497.

Zhang Y. *I-TASSER server for protein 3D structure prediction.* BMC Bioinformatics, 2008. **9**, 40.

Zubrzycka M. and Janecka A. *Substance P: transmitter of nociception (Minireview).* Endocr Regul, 2000. **34**, 195-201.