CHARLES UNIVERSITY IN PRAGUE

FACULTY OF SCIENCE

DEPARTMENT OF BIOCHEMISTRY

STUDIES ON STRUCTURE AND BIOLOGICAL FUNCTIONS OF NKR-P1 RECEPTORS

Ph.D. THESIS

Supervisor: RNDr. Petr Novák, Ph.D.

PRAGUE 2013

DANIEL ROZBESKÝ

To my dearest parents Jaroslava and Daniel Rozbeský

DECLARATION

I declare that I have worked on this thesis under the guidance of my supervisor and that all sources of the previous knowledge are properly cited. No part of this work was used and will not be used for obtaining any other academic degree than Ph.D. from Charles University in Prague.

Prague …….. ……..…..……

Daniel Rozbeský

DECLARATION OF AUTHORSHIP

I declare that Daniel Rozbeský contributed significantly (30-90%) to the experiments and to all 6 scientific publications contained in this Ph.D. thesis. He performed most of the experiments, substantially contributed to their planning, and took a significant part in the primary data interpretation and their preparation for the publication.

Prague …….. ……..…..……

RNDr. Petr Novák, Ph.D.

ACKNOWLEDGEMENTS

My deepest gratefulness goes to my parents for their encouragement and endless support which helped me sail through the first scientific chapters of my life. Next, I would like to thank Miška for her love and tolerance during all of my overnight and weekend experiments.

I would like to express my gratitude to my supervisor, Petr Novák, for his continuous support, guidance and valuable advice. I sincerely appreciate the great freedom in performing experiments and creating a good spirit in the laboratory. During coffee discussions, Petr supplied me with many resources and ideas. He showed me the world of science not only in the proteins and structures, but at number of scientific conferences as well.

A very special thanks goes to Karel Bezouška. In spite of his ticklish situation in 2012 and his activities which are absolutely unacceptable to me, I still consider him to be one of my greatest mentors who showed me the pleasure of scientific discovery.

I would also like to thank Daniel Kavan whose door was always open when I needed advice. Furthermore, I would like to express my thanks for creating a great atmosphere in the lab to my labmates Hynek Mrázek, Ljubina Ivanova, Zdeněk Kukačka, Eliška Pospíšilová as well as to all of my previous and current students, namely Martin Čonka, Franta Vostárek, Gábina Petrovová, Pavla Vaňková, Lucka Hernychová and David Adámek.

I wish to thank Petr Man for introducing me to the H/D exchange and for his valuable comments and editing several manuscripts. I want to also thank Josef Chmelík for NMR measurements, Petr Kolenko and Jan Dohnálek for their work in the field of protein crystallography, Julien Marcoux and Carol Robinson for ion mobility measurements, Vladimír Kopecký and Kateřina Hofbauerová for Raman spectroscopy measurements, Jiří Černý for molecular dynamics simulations and Cheyenne Falat for editing manuscript of this thesis.

Many thanks also go to Žofie Sovová and Rudiger Ettrich for their molecular modeling and for showing me the beauty of molecular dynamics.

Finally, this research would not have been carried out without the financial support of the Charles University Grant Agency (Grant 403211/2011 and Project UNCE 204025/2012) and the Grant Agency of the Czech Republic (Grant P207/10/1040).

ABSTRACT

Natural killer (NK) cells play a significant role in the detection and destruction of virally infected and tumor cells. The NKR-P1 receptors regulate NK cell function by an alternative missing-self recognition system. Although the NKR-P1 receptors were among the first surface NK receptors identified on rodent NK cells more than 20 years ago, there is still very little known about their biological function and their physiological ligands. Furthermore, no three-dimensional structure of any of the NKR-P1 family receptors has been published so far.

To understand the functional architecture of mouse NKR-P1 receptors, we developed a simple and efficient protocol providing large amounts of pure soluble NKR-P1 proteins. The crystal structure of mouse NKR-P1A, determined at 1.7 Å resolution, is the first structure of a representative of the NKR-P1 family. Crystal structure is formed by a compact C-type lectin-like domain and an extended loop that participates in domain swapping. A potential role of the swapped loop has been suggested in natural ligand binding by *in silico* studies. However, chemical cross-linking and H/D exchange in combination with high resolution mass spectrometry revealed this loop in close proximity to the compact core in solution. The discrepancy between the crystal and solution structure indicates a testable hypothesis for the conformational change of the loop induced by ligand binding.

ABBREVIATIONS

ADCC	**A**ntibody **D**ependent **C**ellular **C**ytotoxicity
BSA	**B**ovine **S**erum **A**lbumin
CCS	**C**ollisional **C**ross **S**ection
CD	**C**luster of **D**ifferentiation
CTL	**C**ytotoxic **T** **L**ymphocyt
CTLD	**C**-**T**ype **L**ectin-like **D**omain
DC	**D**endritic **C**ell
DLS	**D**ynamic **L**ight **S**cattering
DMSO	**D**i**M**ethyl **S**ulf**O**xide
DSG	**D**i**S**uccinimidyl **G**lutarate
DSS	**D**i**S**uccinimidyl **S**uberate
DTT	**D**i**T**hio**T**hreitol
EDC	1-**E**thyl-3-(3-**D**imethylaminopropyl)**C**arbodiimide hydrochlorid
EDTA	**E**thylene**D**iamine**T**etraacetic **A**cid
ESI	**E**lectro**S**pray **I**onization
FC	**F**ragment **C**rystalizable
FT-ICR	**F**ourier **T**ransform – **I**on **C**yclotron **R**esonance
H/D	**H**ydrogen/**D**euterium Exchange
HEPES	4-(2-**H**ydroxym**E**thyl)-1-**P**iperazin**E**thane**S**ulfonic acid
HLA	**H**uman **L**eukocyte **A**ntigen
HPLC	**H**igh **P**erformance **L**iquid **C**hromatography
HSQC	**H**eteronuclear **S**ingle **Q**uantum **C**oherence
IFN	**I**nter**F**ero**N**
IgSF	**I**mmuno**g**lobulin **S**uper**F**amily
IL	**I**nter**L**eukin
IPTG	**I**so**P**ropyl-β-D-**T**hio**G**alaktopyranosid
ITAM	**I**mmunoreceptor **T**yrosin-based **A**ctivation **M**otif
ITIM	**I**mmunoreceptor **T**yrosin-based **I**hibition **M**otif
KIR	**K**iller cell **I**mmunoglobulin-like **R**eceptors
LB	**L**uria-**B**ertani medium

LRC	**L**eukocyte **R**eceptor **C**omplex
MALDI	**M**atrix **A**ssisted **L**aser **D**esorption **I**onization
MCMV	**M**ouse **C**yto**M**egalo**V**irus
MHC	**M**ajor **H**istocompatibility **C**omplex
MS	**M**ass **S**pectrometry
NK	**N**atural **K**iller
NKC	**N**atural **K**iller **C**omplex
NKR-P1	**N**atural **K**iller **R**eceptor – **P**rotein **1**
OD	**O**ptical **D**ensity
PAMP	**P**athogen-**A**ssociated **M**olecular **P**attern
PCR	**P**olymerase **C**hain **R**eaction
PEG	**P**oly**E**thylene **G**lycol
PIPES	**PIP**erazine-N,N′-bis(2-**E**thane**S**ulfonic acid)
PMSF	**P**henyl**M**ethyl**S**ulfonyl **F**luoride
PVDF	**P**oly**V**inyli**D**ene **F**luoride
RCMV	**R**at **C**yto**M**egalo**V**irus
SDS	**S**odium **D**odecyl **S**ulfate
TCA	**T**ri**C**hloroacetic **A**cid
TCEP	**T**ris(2-**C**arboxy**E**thyl)**P**hosphine
TCR	**T**-**C**ell **R**eceptor
TFA	**T**ri**F**luoroacetic **A**cid
TGF	**T**ransforming **G**rowth **F**actor
T_H	**H**elper **T** Cell
TLR	**T**oll-**L**ike **R**eceptors
TNF	**T**umor **N**ecrosis **F**actor
TOF	**T**ime **O**f **F**light
Tris	**TRIS**(hydroxymethyl)aminomethane

TABLE OF CONTENTS

CHAPTER 1
INTRODUCTION

1.1 NK cell mediated immunity

Natural killer (NK) cells represent an important component of innate immunity that play a key role against virally infected and tumor cells. They were originally identified more than 35 years ago through their capacity to kill certain tumor cells *in vitro* in the absence of previous stimulation[1]. Due to the expression of several lymphoid markers as well as their origin from the lymphoid progenitor, NK cells are regarded as lymphocytes. They do not express any antigen-specific receptors, thus they are considered to be components of innate immunity. On the other hand, many authors propose their placement as an evolutionary bridge between innate and adaptive immunity because they display several features of adaptive immunity such as memory of previous exposure to a specific antigen[2,3]. Unlike T and B lymphocytes, NK cells mediate their effects through a variety of cell surface inhibitory and activating receptors. Moreover, NK cells produce several classes of immunoregulatory cytokines, which help shape immune response. Orchestrated interactions between these molecules may finally lead to the killing of target cells[4,5].

One of the most discussed topics in the field of NK cells is their role in antitumor immunity, which was documented in many experimental models. For example, mice depleted of NK cells were observed to be more susceptible to spontaneous or induced tumors[6]. In humans, although selective NK deficiencies are extremely rare, epidemiological studies have revealed a similar tight correlation between low NK cell activity in peripheral blood and increased cancer risk[7]. Furhermore, NK cell infiltration into tumors has been shown to be associated with a better prognosis in gastric and colorectal carcinoma[8], renal carcinomas[9] and squamous cell lung carcinomas[10,11].

Functional and structural studies of particular NK receptors shed light on molecular mechanisms of target cell recognition. One of the most explored NK receptor in antitumor immunity is the activating receptor NKG2D, which has been demonstrated to play an important role in the control of tumor growth in several experimentally induced

tumors. NKG2D binds the ligands RAE-1, H60, or MICA/B, which are surprisingly over-expressed in several tumor cell lines. It was shown that engagement of NKG2D with ligands results in apoptosis of the target cell[12]. The other activating receptor NKp30 has also been observed in tumor cell recognition and lysis, particularly of neuroblastoma and glioblastoma cell lines[13]. Moreover, some data indicate its role in acute myeloid leukemia because the loss of NKp30 expression in patients with this disease is associated with impaired NK cell cytotoxicity and shorter survival[14].

Recent advances in molecular mechanisms of NK cell biology are being harnessed to develop NK cell-directed immunotherapies for cancers and autoimmune diseases. Potential promising strategy is manipulating between positive and negative signals provided by activating and inhibitory receptors, respectively. One of these approaches, which has been shown in tumor rejection in preclinical models of leukemia and melanoma[15], is based on specific antibodies that block inhibitory receptors. Further strategy exerts bispecific antibodies in an effort to increase the efficiency of antibody-dependent cellular cytotoxicity (ADCC) at tumor sites. The bispecific antibody usually combines specificity for the tumor antigen such as HER2/neu and the FcγRIII receptor (CD16) found on NK cells[16]. Nevertheless, the most hotly debated strategy in the past decade is based on NK cell alloreactivity and graft-versus-leukemia effect in the acute myeloid leukemia treatment. In this approach, improved disease-free survival and reduced graft-versus-host disease (GVHD) was observed in patients after haploidentical stem cell transfer[17-19], although some discrepancies occurred in subsequent studies[20].

As described previously, NK cells play a significant role against viruses as well. The NK cell-mediated antiviral immunity has been observed in many studies showing increased susceptibility to viruses in mice depleted of NK cells or in mouse models with defects in cytotoxicity as well as in human NK deficiencies[21-24]. NK cells mediate their antiviral effect through two different mechanisms: cytokine release and induction of apoptosis in target cells through cell surface receptors. For example, in mice infected with mouse cytomegalovirus (MCMV), the surface molecule m157, encoded by MCMV, is selectively recognized by activating NK receptor Ly49H which leads to host protection[25,26]. Further evidence for the importance of NK cells during viral infections has been demonstrated by expression of multiple virally encoded decoy ligands for inhibitory NK receptors in virally infected cells[27,28].

Besides antitumor and antiviral activity, NK cells play a crucial role as regulatory cells by directly acting on DCs, T cells, B cells, and endothelial cells. Two different effects have been observed in the case of DCs: contact-dependent killing of immature DC by NK cells and facilitated cross-presentation of antigen from apoptic NK cells targets[29]. In case of T cells, NK cells can participate in priming $CD4^+$ T cells by producing IFN-γ[30,31], while in case of B cells, NK cells can suppress the B cell autoreactivity[32]. Some data indicate that NK cells may be involved in the killing of endothelial cells during viral infection and thus in the pathogenesis of vascular injury[33,34]. On the other hand, a very distinct subset of NK cells, known as uterine NK (uNK) cells, secrete several proangiogenic factors such as placental growth factor or vascular endothelial growth factor, both of which are crucial for placental development[35,36].

In conclusion, numerous experimental and epidemiologic data support the significant role of NK cells in antiviral defense, antitumor immunity, hematopoietic stem cell transplantation, immunoregulation, and reproduction. However, many aspects of NK cell functions remain yet to unveiled[37].

1.2 NK cell recognition

Specificity of NK cells for target cells is determined by a sophisticated repertoire of activating and inhibitory receptors expressed on the surface of NK cells. Activation of NK cells is dependent not only on the amount of activating and inhibitory ligands on the target cell, but also on the presence of specific NK receptors for these ligands and, finally, on the balance between activating and inhibitory signals transduced through NK receptors. So far, three hypotheses have been proposed to explain NK cell recognition.

The first concept, based on the capacity of NK cells to attack target cells that fail to express sufficient levels of MHC class I molecules, is referred to as missing-self recognition (**Fig. 1**). This concept was originally explained by the expression of inhibitory NK receptors specific for MHC class I molecules. NK cells were proposed to attack any cell that had down-regulated expression of MHC class I molecules because the target cell could not engage with the inhibitory NK cell receptor for MHC class I. In addition to this, the down-regulation of MHC class I is a common mechanism of virally infected and tumor cells to evade cytolysis by $CD8^+$ T cells. Therefore the missing-self recognition by NK cells seems to be a counter-response by the immune system[38,39].

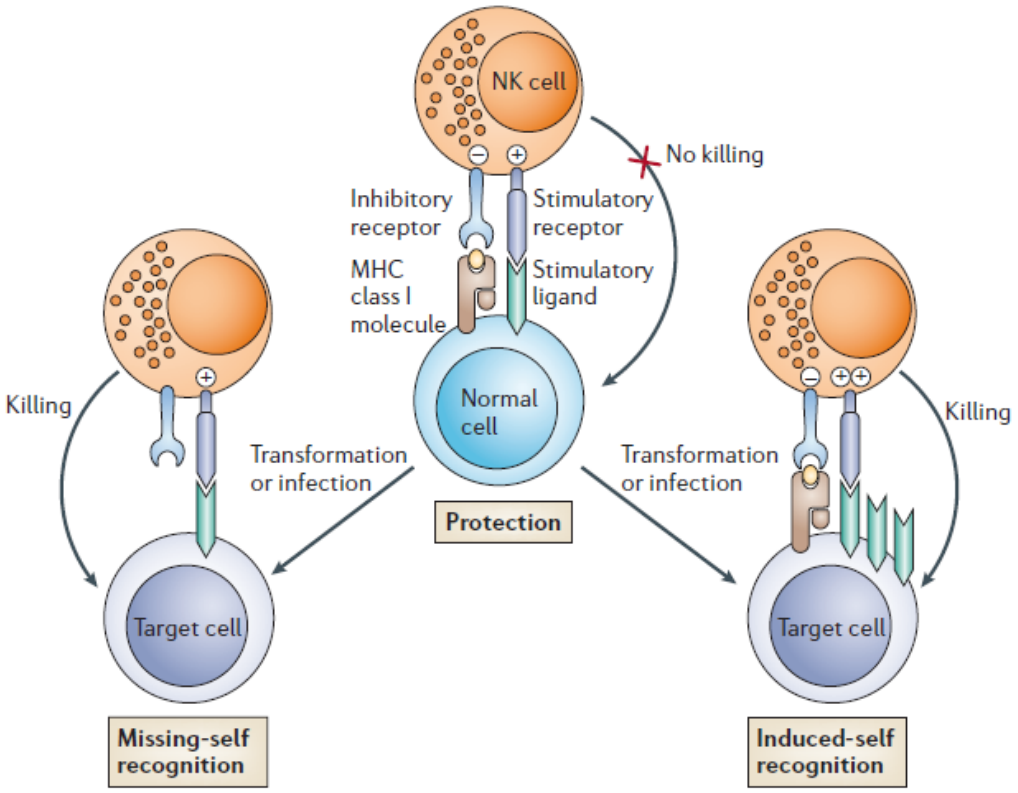


Figure 1 *Activity of NK cells is determined by the dynamic equilibrium between activating and inhibitory signals transduced through NK receptors. NK cells are tolerant to normal cells because the strength of the activating signals is dampened by engagement of the inhibitory receptors. Absence of MHC class I molecules or over-expression of activating ligands in target cell result to overcome the inhibitory signals which finally lead to NK cell activation*[57]*.*

Recent studies demonstrate that while the engagement between inhibitory receptors and MHC class I molecules is essential for target cell recognition, down-regulation of MHC class I molecules on target cell is not sufficient to fully trigger NK cell-mediated cytotoxicity. Furthermore, sufficiently potent activating NK receptors must be stimulated[40] by ligands which are expressed by even normal cell types[41]. This contemporary modification of the missing-self concept can explain the inability of NK cells to attack cells with no (e.g., erythrocytes) or low (e.g., neural cells) expression of MHC class I molecules[42]. Nevertheless, the other explanation of this phenomenon relies on recognition of non-MHC ligands, which has been observed by interaction of the mouse inhibitory receptor NKR-P1B and its broadly expressed ligand Clr-b[43].

Besides the missing-self recognition via inhibitory receptors, NK cells have been shown to detect internal changes of self molecules in damaged tissues using activating receptors. This concept referred as stress-induced self recognition is based on the recognition of self molecules whose expression is usually barely detectable in steady-state conditions; however, it is up-regulated during various forms of stress[44-46]. Prototypical examples of this concept are recognition of B7 proteins expressed in tumor cells by activating receptor NKp30,[47,48] or recognition of stress-inducible proteins such as MICA or MICB by activating receptor NKG2D[44,49].

Further mechanism of NK cell recognition is antibody-dependent cellular cytotoxicity (ADCC). This process relies on the capacity of NK cells to detect antibody coated cells through the low affinity FcγRIIIA surface receptor for IgG also known as CD16[50]. This ADCC mechanism has been observed in several monoclonal antibody-based therapies, including treatment with rituximab[51,52].

1.3 NK cell education and tolerance

An important feature of immune system is the capacity to discriminate between self and non-self. Like T cells and B cells, NK cells acquire tolerance to self in order to avoid destruction of normal tissue. The requirement for self tolerance of NK cells is particularly a result of the large diversity of NK receptors and their ligands. Although more than 20 inhibitory receptors are known, each NK cell expresses an average of 3-5 different inhibitory receptors[53,54]. Each inhibitory receptor is not expressed by all NK cells, but only by a subset of NK cells, overlapping in expression with other family members[55]. Both NK receptors and MHC class I molecules exhibit high allelic polymorphism, and each inhibitory receptor binds only certain MHC class I variants but not others. Moreover, the genes for NK receptors and MHC class I molecules lie on different chromosomes and are inherited independently, thus genetic mechanism is not able to secure that NK receptors of an individual exhibit specificity for the appropriate self MHC class I molecule[56]. To explain NK self-tolerance, several hypotheses have been proposed recently[57].

An early promising model known as at-least-one suggested that all NK cells express at least one inhibitory receptor specific for self MHC class I, however, this fact has never been proven[55,58]. Today, a widely accepted model proposes that NK cells participate in a

unique process referred to as NK cell education. This process either leads to the maturation of the functional NK cell repertoire or to the induction of a hyporesponsive state in potentially autoreactive NK cells, thereby become tolerant to self. Indeed, a fraction of hyporesponsive NK cells lacking inhibitory receptors for self MHC I was observed in mice[59] or humans[60]. Further support for the latter mechanism came from experiments with NK cells in MHC I deficient individuals in which NK cells should be chronically activated but instead are hyporesponsive to self cells as well as to other stimuli[61-63]. Taken together, two mechanisms of NK self-tolerance are known: first, the effector functions of functionally competent NK cells are inhibited by the recognition of self MHC I, and second, hyporesponsive NK cells coexist that are not able to detect self MHC I[64]. It is noteworthy that the switch from a competent state to hyporesponsive is not induced by changes in the transcriptional program, rather seems to be based on the confinement of activating receptors in membrane nanodomains[65-67]. Despite of significant progress in research of NK self tolerance, many molecular mechanisms are still unclear.

1.4 NK receptor signal transduction

NK cells share many common features of signaling pathways with the immunoreceptors expressed on B cells and T cells. In general, signal transduction of inhibitory NK receptors (**Fig. 2**) is mediated through one or more immunoreceptor tyrosine-based inhibitory motifs (ITIM) in their cytoplasmic domain with the consensus sequence Ile/Val/Leu/Ser-x-Tyr-x-x-Leu/Val, where x denotes any residue[68].

Upon ligand binding, the tyrosine residues in the ITIM are phosphorylated by a Src family tyrosine kinases, which then recruit and activate protein-tyrosine phosphatases such as SHP1, SHP2 or SHIP1 depending on particular receptor. Activation of phosphatases results in decreased phosphorylation of numerous signaling proteins, including PLCγ1, PLCγ2, ZAP70, Syk, Shc, LAT, Vav-1, FcεRIγ or SLP76, and subsequent interruption of the early phosphorylation pathways that are responsible for NK cell activation[69-73].

Many activating NK receptors lack the cytoplasmic domain with ITIMs. Instead, they associate with adaptor proteins such as CD3ζ[74], FcεRIγ[75] or DAP12[76] bearing immunoreceptor tyrosine-based activation motifs (ITAM) in the cytoplasmic domain

with the conserved sequence Asp/Glu-x-x-Tyr-x-x-Leu/Ile-x_{6-8}-Tyr-x-x-Leu/Ile[77]. Association between activating NK receptor and ITAM adaptor is mediated by electrostatic interaction of oppositely charged residues within transmembrane segments. Engagement of ITAM adaptors activates tyrosine kinases of the Src family such as p56lck and p59fyn, which phosphorylate tyrosine residues on ITAMs, thereby leading to the recruitment and activation of tyrosine kinases of the Syk family such as Syk and ZAP70. Several transmembrane and cytoplasmic proteins then enter into signaling pathways, which finally results in Ca^{2+} influx, degranulation, and transcription of cytokine genes[78].

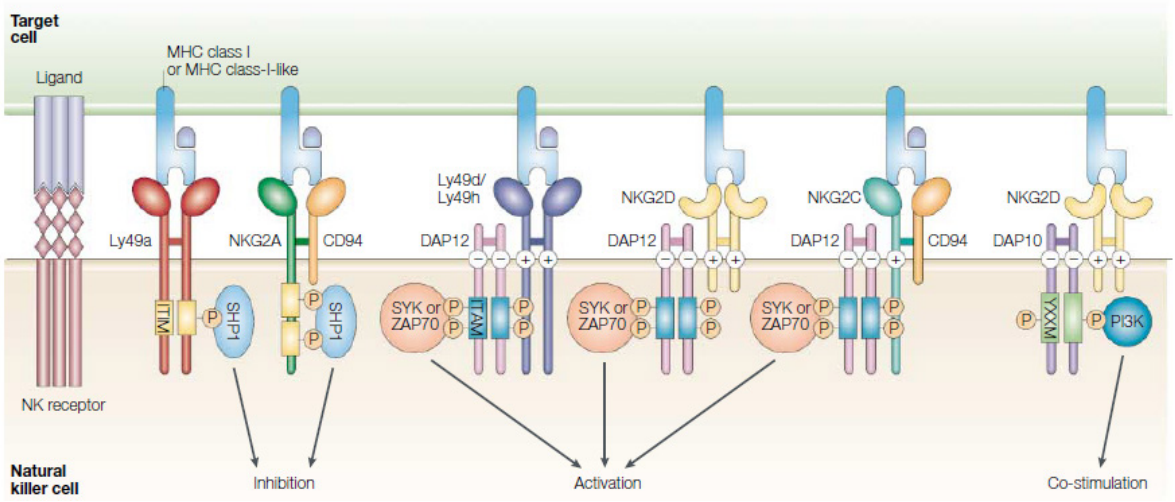


Figure 2 *NK cell effector signaling pathways. Upon interaction of inhibitory receptors with ligand, the cytoplasmic domains become tyrosine phosphorylated on ITIMs and associate with intracellular phosphatases, including SHP1, SHP2 or SHIP1, resulting in an inhibitory signal. The activating receptors are associated with ITAM bearing adaptors via electrostatic interactions. Upon ligand binding, ITAMs are tyrosine phosphorylated, resulting in Ca^{2+} influx, degranulation and transcription of cytokine genes[56].*

In the case of activating receptor NKG2D, several different features in signaling pathways have been observed. NKG2D exists in two alternatively spliced isoforms: the long (NKG2D-L) and short (NKG2D-S). While NKG2D-L associates exclusively with the other adaptor protein DAP10, NKG2D-S is able to associate with both DAP10 and DAP12. In addition to this, DAP10 is a transmembrane adaptor protein with the short cytoplasmic domain containing the conserved motif Tyr-x-x-Met that binds phosphatidylinositol-3 kinase and Grb2 upon tyrosine phosphorylation. Thus, in activated NK cells the NKG2D is able to induce both ITAM-based and

phosphatidylinositol-3 kinase associated pathways[79,80].Finally, the effector functions of NK cells depend on a dynamic equilibrium between multiple and sometimes opposing signals transduced by NK receptors.

1.5 Structure and function of NK receptors

1.5.1 General features of NK receptors

The most important NK receptors (**Fig. 3**) belong to the two major classes, those of the immunoglobulin superfamily (IgSF) and of the C-type lectin-like superfamily. The immunoglobulin superfamily consists of the killer-cell immunoglobulin-like receptors (KIR) and the leukocyte immunoglobulin-like receptors (LIR). The C-type lectin-like superfamily contains the Ly49 family, CD94/NKG2 family, NKR-P1 family, Clr family and many other members, including CD69, AICL, LLT1 and MAFA.

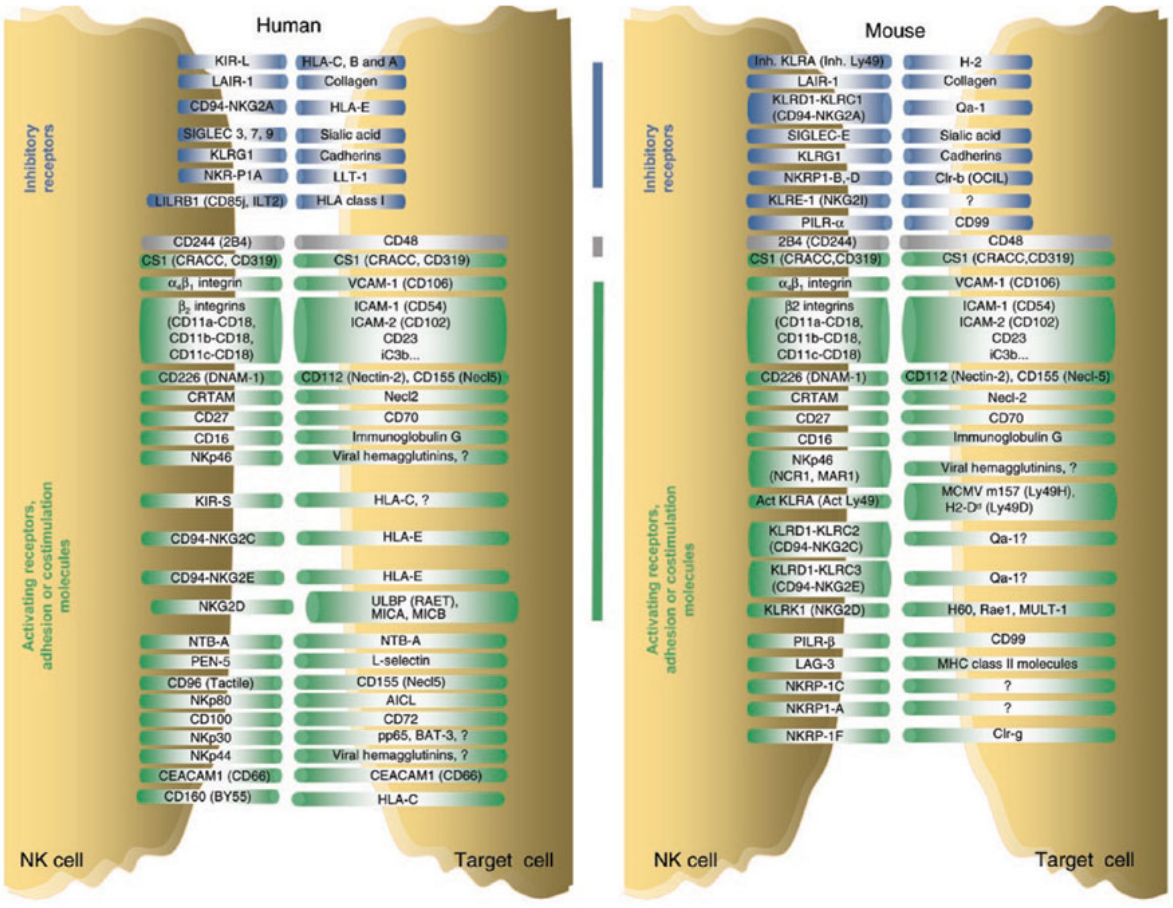


Figure 3 *Prominent human (left) and mouse (right) NK receptors in interaction with their ligands. Inhibitory receptors are in blue; 2B4 molecule which can act as an activating or an inhibitory receptor is in gray; the other receptors are in green[37].*

In addition, the classification into families is based on a high level of homology, typically more than 80%. Between families, the receptors share some structural features; however they are distantly related in sequence[81].

From a structural point of view, the IgSF receptors are type I (extracellular N terminus) transmembrane proteins consisting of either two or three extracellular immunoglobulin domains, whereas the C-type lectin-like receptors are type II (extracellular C terminus) trasmembrane proteins containing a single extracellular C-type lectin-like domain. Although these receptor families are structurally different, they share some key characteristics, such as signaling pathways or specificity for various MHC class I molecules. Interestingly, both superfamilies include activating as well as inhibitory receptors. Moreover, some of the NK receptors are broadly expressed on a broad range of cell types, mostly of the haematopoietic lineage[42,81].

The genes of IgSF are located together in the leukocyte receptor cluster (LRC) on human chromosome19q13.42[82], while the genes of C-type lectin-like receptors are encoded by the natural killer gene complex (NKC) on the human chromosome 12p12-13, mouse chromosome 6, and rat chromosome 4[83,84]. The genes in both gene clusters have exon/intron organization, and the genes coding activating receptors usually have a mutation in the cytoplasmic domain, which results in a stop codon and loss of the ITIM motif[82,85]. This fact, in combination with phylogenetic analysis, suggests that activating receptors have evolved rather recently from ancestral inhibitory receptors due to selection for resistance to infections. This is probably the reason for difficult identification of their ligands and their role in NK mediated immunity[86,87].

A remarkable relationship has been found between KIR and Ly49 genes. Whereas the Ly49 locus has rapidly expanded in mice (23 genes) and rats (26 genes), only three KIR genes are found in these rodents. Alternatively, the KIR locus has expanded in primates, but efforts to identify human Ly49 genes have detected only a single Ly49L gene that is probably not functional due to an early stop codon in an extended exon 5. This fact, together with comparison of KIR and Ly49 haplotypes in other species, suggests rapid evolutionary change since the divergence of a common ancenstor[88-90]. Therefore surprisingly, the human KIR and their rodent analogs Ly49 do not bind their MHC class I ligand in the same manner but they interact at a different sites[81].

Both KIR and Ly49, as well as some other NK receptors, exhibit a high level of diversity. Indeed, many NK receptors exist in different isoforms which differ from each other in point mutations, small deletions or insertions, or even chimeric genes[91]. The reason for this significant repertoire diversity of NK receptors and their rapid evolution is the co-evolution with MHC genes, which are one of the most polymorphic molecules encoded in mammalian genome. Thus changes in MHC class I impose pressure on the evolution of NK receptors in order to maintain specificity with their MHC I ligands[92]. In this interesting example of convergent evolution of the NK receptors, the NK genes expansion and diversification was accelerated especially by the gene duplication and gene conversion which was probably facilitated by their close proximity and tandem head-tail arrangement[82,93].

Furthermore, monoallelic expression and promiscuity in MHC class I recognition, which was observed for the Ly49 as well as KIR, also significantly contributes to the repertoire diversification and complexity[94]. Finally, deeper understanding of repertoire diversity and genetic variability is a prerequisite for sophisticated modulation of NK cell functions for therapeutic reasons.

1.5.2 Killer cell immunoglobulin-like receptors

The human killer immunoglobulin-like receptors (KIR) represent a family of type I transmembrane proteins expressed on the surface of NK cells and a subset of T cells. The KIR family is encoded by 15 highly polymorphic genes and 2 pseudogenes. Terminology of KIR is based on the number of immunoglobulin-like domains in the extracellular part (2D for two domains, 3D for three domains) and on the length of the cytoplasmic domain (L for long domains, S for short ones)[95]. All inhibitory KIR possess long cytoplasmic domains, while all activating receptors have short cytoplasmic domains with the exception of 2DL4[96]. In addition to this, the immunoglobulin-like domains are designated D0, D1, and D2.

The ligands for KIR are mainly HLA molecules. Interestingly, activating KIR also recognize HLA molecules, however, their affinity is lower than inhibitory KIR[97]. On the other hand, activating KIR seem to respond better to allogenic HLA which provides support for the concept by which NK cells promote anti-tumor response after haematopoietic transplantation[98].

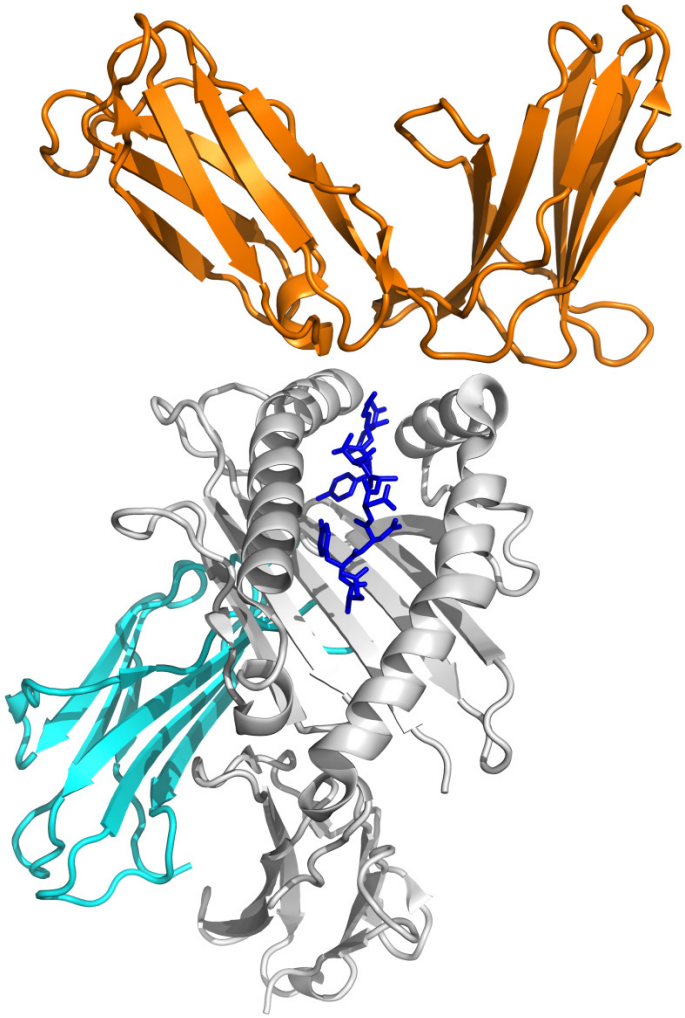


Figure 4 *Ribbon representation of the 2DL1-HLA-Cw4 complex colored by chains: 2DL1, orange; HLA-Cw4, gray; ß2m, cyan; peptide, blue. PDB code:1IM9*

The structure of the extracellular region of KIR2D consists of two immunoglobulin-like domains, D1 and D2, which are connected by a linker of three to five residues. Each domain is formed by two antiparallel ß sheets which pack to each other. Both domains have V-shape orientation and the angle between domains varies significantly in individual KIR2D[99-101]. Some authors suggest that angle alterations are responsible for varying affinity to HLA molecules[102].

Whereas TCR/HLA interaction is based especially on hydrogen bonds and van der Waals forces, the KIR/HLA binding relies on hydrogen bonds and electrostatic interactions. In the 2DL1/HLA-C complexes (**Fig. 4**), the D1 and D2 domains bind their HLA-C ligand through α1 and α2 helices, and C-terminal part of bound peptide[103]. The interaction buries approximately 1500 Å^2 of solvent accessible area, which is comparable to the buried surface of TCR/HLA complexes[104,105]. At least 16 conserved residues participate in contact between HLA-C and 2DL2 via salt bridges and hydrogen bonds. The site directed mutagenesis of these residues revealed that mutation of each residue to alanine resulted to a 20 fold lower affinity[103].

Allelic specificity of 2DL receptors for HLA-C allotypes is based on the polymorphism at positions 77 and 80 of the HLA-C heavy chain. 2DL1 receptors have C2 specificity and

are able to recognize HLA-C alleles with Asn77 and Lys80 which are present in HLA-Cw2, HLA-Cw4, HLA-Cw5, and HLA-Cw6, while 2DL2 receptors exhibit C1 specificity and recognize HLA-C with Ser77 and Asn80 (HLA-Cw1, HLA-Cw3, HLA-Cw7, HLA-Cw8)[100].

In the case of peptide selectivity, KIR recognition is both peptide dependent and peptide selective. However, these receptors are not able to distinguish between self and non-self peptide[106-109].

In spite of the fact that the structure of KIR3D has not been determined so far, recent analyses indicate that the D0 domain participates in direct contact with HLA[110].

1.5.3 The Ly49 family

The Ly49 is a multigene family that encodes homodimeric type II transmembrane proteins with a single C-type lectin-like domain. Ly49 receptors represent the major functional NK receptor family in mice with expression on NK cells, NKT cells and also on a subset of $CD8^+$ T cells. Like their human analogs KIR, the Ly49 receptors mainly recognize MHC class I molecules, H-2K and H-2D. The exact number of mouse Ly49 loci has not been determined, however, there are at least 23 genes (Ly49A-W). Moreover, the Ly49 genes exhibit high level of polymorphism and the number of loci is dependent on the mouse strain[111-113].

From a structural point of view, Ly49 receptors encompass the C-type lectin-like domain, which is connected to the transmembrane segment using a stalk region. The intracellular part consists of a short or a long cytoplasmic domain depending on activating or inhibitory function, respectively. The C-type lectin-like domain, found in other members of the C-type lectin-like family as well, includes two antiparallel ß-sheets and two α-helices. Six cysteines, which are the most conserved residues in the C-type lectin-like domain, form disulfide bonds with conserved arrangement C1-C2, C3-C6, and C4-C5[114].

In the crystal structure of complex Ly49A-$H2D^d$ (**Fig. 5a**), the Ly49A homodimer binds $H2D^d$ asymmetrically, with only one of its two subunits binding a single MHC I molecule. The contact site is at a broad cavity underneath the peptide binding platform of MHC and partially overlaps the CD8 binding site[115]. On the other hand, in the Ly49C-$H2K^b$ complex (**Fig. 5b**), the Ly49C homodimer contacts $H2K^b$ bivalently at a site overlapping the Ly49A binding site on $H2D^d$ to form a symmetrical butterfly-shaped

assembly[116]. The geometry of Ly49A and Ly49C in the complexes described above is slightly different. While the Ly49C homodimer adopts an open conformation, the Ly49A homodimer adopts a closed conformation, in which C-type lectin-like domains are more closely juxtaposed than in Ly49C. However, NMR structure of Ly49A unveiled that in solution, the protein exists predominantly in the open conformation, allowing it to bind two MHC class I molecules[117].

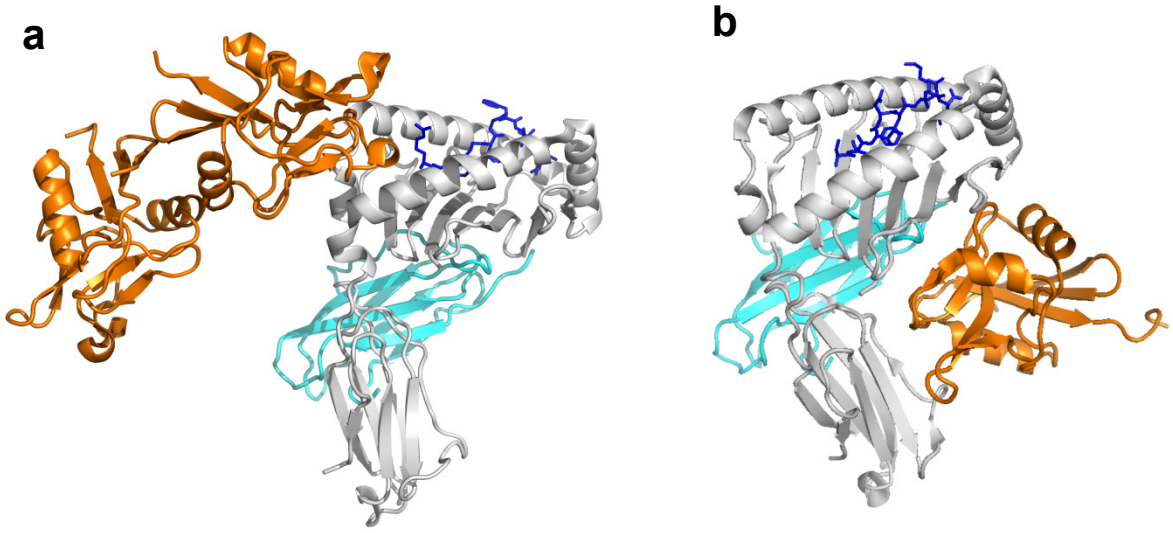


Figure 5 *Ribbon representation of the Ly49A-H2D^d (**a**) and Ly49C-H2K^b (**b**) complexes colored by chains: Ly49, orange; H2, gray; ß2m, cyan; peptide, blue. PDB codes:1QO3 and 3C8K*

Different modes of MHC class I engagement by Ly49A and Ly49C raised a question of the *cis-trans* interaction mediated by different Ly49 conformations. In *trans* interaction the cell surface receptor interacts with ligands expressed on the other cells, while in *cis* interaction the receptor engages ligands on the same cell[118]. The *cis* interaction has been observed for some members of Ly49 and LILRB or PIR receptor families[119-121]. The role of *cis* interaction in NK cell biology seems to be involved in NK cell education and in a modulation of inhibitory capacity of Ly49[122,123]. The structural basis for *cis-trans* binding by Ly49 using the back-folded and extended conformation has been recently observed[124].

Similarly to KIR, interaction between Ly49 and MHC class I molecules is dependent on the presence of a bound peptide within the H2 groove. In case of Ly49A, interaction is independent of peptide sequence. On the other hand, low degree of peptide selectivity has been observed for Ly49C and Ly49I[125,126]. Although the Ly49 evolved from C-type lectin family, specificity of Ly49 to carbohydrates or Ca^{2+} has not been observed.

While the function of inhibitory Ly49 receptors is monitoring the expression level of MHC class I molecules, the role of many activating Ly49 receptors remains an enigma. Though some activating receptors such as Ly49P and Ly49W have also been reported to bind $H2D^d$, their function has not been investigated[127,128]. Surprisingly, activating receptor Ly49H exhibits no binding to any known H2 molecule, but it binds to the m157 viral surface glycoprotein encoded by MCMV. This observation suggests the hypothesis of non-self protein recognition by activating Ly49 receptors[25,26].

1.5.4 The CD94/NKG2 family

The CD94/NKG2 family consists of disulfide-linked heterodimers expressed as type II transmembrane proteins on human and rodent NK cells and a subset of T cells[129,130]. While the CD94 subunit is invariant, the NKG2 subunit comprises at least five proteins designated NKG2A, -B, -C, -E and –H. In addition, NKG2A/B and NKG2E/H are splice variants[131]. The NKG2D is a homodimer and its structure and function is discussed separately bellow. The cytoplasmic domain is either short (NKG2C and –E/H) or long (NKG2A/B) corresponding to their activating or inhibitory function, respectively[76,132-134].

The NKG2 subunit of CD94/NKG2 contains C-type lectin-like domain which binds the non-classical MHC class Ib molecules HLA-E[135-137] and Qa-1[138], in humans and mice, respectively. HLA-E and Qa-1 present peptides derived from the relatively conserved leader sequence of classical MHC class I molecules and HLA-G, and thus serve as indicators for normal expression of MHC class I molecules[139-141]. In comparison to the KIR and Ly49, the CD94/NKG2 receptors exhibit low level of polymorphism, probably due to very little allelic variations in their ligands.

In the crystal structures determined recently, the fold of NKG2 is structurally very homologous to the C-type lectin-like domains, while the structure of CD94 revealed a

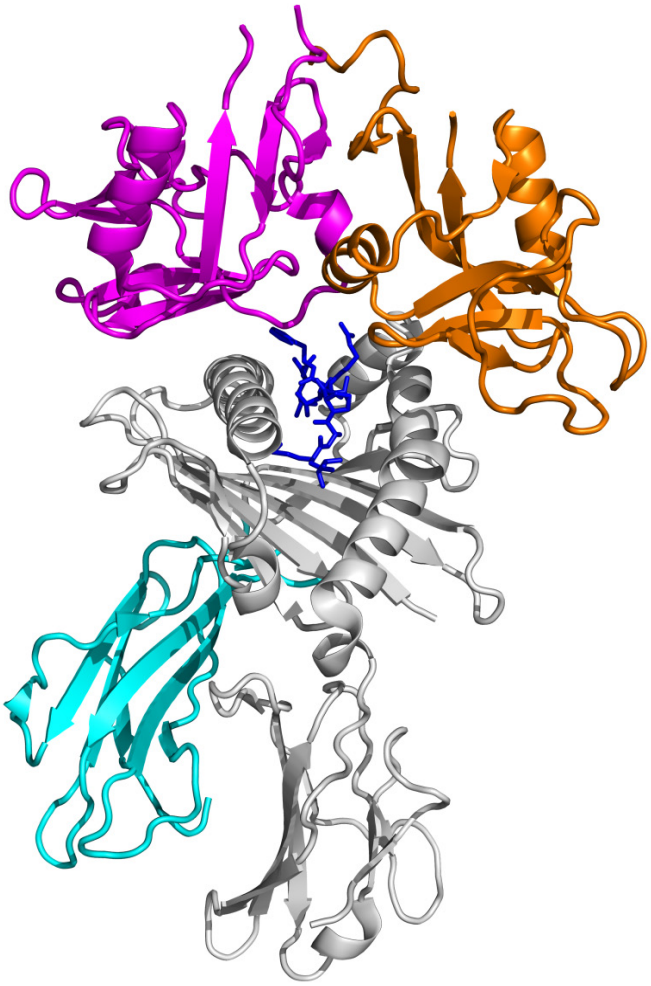


Figure 6 *Ribbon representation of the CD94/NKG2A-HLA-E complex colored by chains: NKG2A, orange; CD94, magenta; HLA-E, gray; ß2m, cyan; peptide, blue. PDB code: 3CDG*

unique variation of C-type lectin fold, in which the second α-helix is replaced by a loop. The structure of the CD94/NKG2A heterodimer showed that dimerization is based not only on an intermolecular disulfide bond in the stalk region but also on asymmetric interactions[142,143].

In the crystal structure CD94-NKG2A in complex with HLA-E bound to a peptide (**Fig. 6**), the CD94/NKG2A subunits lay across the peptide binding groove. In the interface, the CD94/NKG2A subunits interact almost exclusively with α1 and α2 helices of HLA-E via a large number of polar interactions, including salt bridges and hydrogen bonds. Moreover, the CD94 subunits dominate the peptide mediated contacts (80%) compared to NKG2A (20%)[144,145].

1.5.5 The NKG2D receptor

The NKG2D is an activating homodimeric receptor expressed on most NK cells, $CD8^+$ αß T cells and γδ T cells. NKG2D is not a member of CD94/NKG2 family, due to limited sequence similarity to the other members of NKG2 (20-30% sequence identity)[44,146]. Ligands of human NKG2D include MICA and MICB, which are MHC class I homologs composed of α1, α2, and α3 domains. In addition to this, MICA and MICB lack ß2m and do not bind peptides. Further ligands of human NKG2D are the human cytomegalovirus UL16 binding proteins known as ULBP. Unlike MHC class I molecules, expression of

MICA and MICB is up-regulated by cellular stress especially on epithelium-derived tumors[49,147-150]. Murine NKG2D recognizes the retinoic acid-inducible RAE-1 family and the H60 minor histocompatibility antigen. From a structural point of view, both RAE-1 and H60 are homologous to the structure of MHC class I but they lack an α3 domain and are anchored in the membrane by GPI linkage[151,152]. Interestingly, it was observed that some tumors secrete or release soluble NKG2D ligands, which serve as decoy ligands in order to subvert immune response[153-155]. In the crystal structure of the NKG2D-MICA complex (**Fig. 7**), the overall fold of NKG2D is similar to the C-type lectin-like domain of CD94, retaining only one of the α-helices.

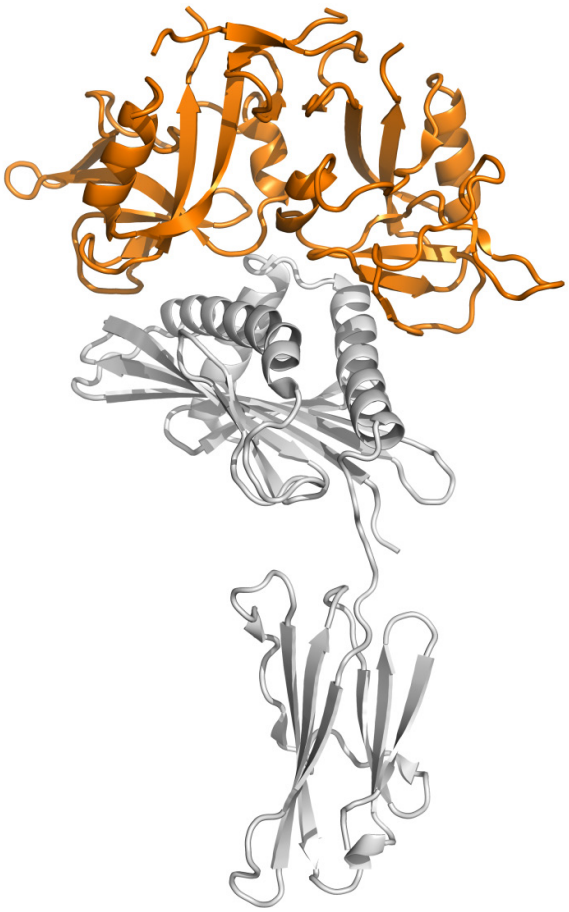


Figure 7 *Ribbon representation of the NKG2D-MICA complex colored by chains: NKG2D, orange; MICA, gray. PDB code:1HYR*

The NKG2D homodimer diagonally overlays the surface of the MICA platform in a manner resembling the docking mode of TCR onto MHC class I. In the NKG2D-MICA interface, each NKG2D monomer mainly contacts either α1 or α2 domains of MICA through a mixture of polar, hydrophobic and ionic interactions. The interaction buries approximately 2180 $Å^2$ of surface area, which is 20-25% greater than TCR-MHC I. Furthermore, the NKG2D-MICA interface is characterized by a high degree of shape complementarity[156]. In the crystal structure of mouse NKG2D-RAE-1, the NKG2D homodimer binds to platform RAE-1 in a manner resembling the interaction of human NKG2D and MICA with the exception of a conformational change in the α2 helix of MICA that undergoes a large structural rearrangement upon receptor binding[157].

1.5.6 The NKR-P1 family

NKR-P1 receptors are type II transmembrane C-type lectin-like receptors expressed on the surface of NK cells and a subset of T cells. The mouse NKR-P1 family consists of seven members, NKR-P1A (encoded by *Klrb1a*), NKR-P1B (encoded by *Klrb1b*), NKR-P1C (encoded by *Klrb1c*), NKR-P1D (encoded by *Klrb1b*), NKR-P1F (encoded by *Klrb1f*), and NKR-P1G (encoded by *Klrb1g*). Whereas NKR-P1D represents an allelic form of NKR-P1B, *NKR-P1E* (*Klrb1-ps*) seems to be a pseudogene[158-160]. In rats, four members of NKR-P1 family have been identified (NKR-P1A, -B, -F and-G)[161], while in humans only a single molecule NKR-P1A (encoded by *KLRB1* or *CD161*) is known[162,163].

It is noteworthy that significant variations in expression of individual NKR-P1 receptors on NK cells have been documented in mice. While NKR-P1C is expressed at high levels on all NK and NKT cells, NKR-P1A and NKR-P1F are expressed at low levels on all NK cells, but not on NKT cells. On the other hand, NKR-P1B is stochastically expressed on approximately 60% of NK cells[43,164,165].

Mouse NKR-P1B(D) contains ITIM in cytoplasmic domains, indicating their inhibitory function, whereas mouse NKR-P1A, -C, -F lack ITIM and have a charged residue in the transmembrane segment which suggests their activating function. Although mouse NKR-P1G is an ITIM bearing receptor, it has been identified only at cDNA level[159]. Association of an ITAM containing adaptor protein FcεRIγ with mouse NKR-P1C has been observed upon antibody cross-linking, which finally led to the induction of NK cell mediated cytotoxicity and IFN-γ production[166,167]. On the other hand, antibody cross-linking of NKR-P1A and NKR-P1F resulted in neither cytokine production nor induction of cytotoxicity[164].

The mouse NKR-P1C corresponds to the NK1.1 alloantigen and represents one of the most important surface markers detected by monoclonal antibody PK136, known also as antiNK1.1[168,169]. Although mouse NKR-P1C has been originally cloned as the NK1.1 antigen, analysis revealed that NK1.1 epitope is shared by two receptors, NKR-P1C and NKR-P1B[170]. The PK136 antibody defines NK cells from CE, B6, NZB, C58, Ma/My, ST, SJL, FWB and Swiss outbred mice, but not from BALB/c, AKR, CBA, C3H, DBA, or 129 mice[168-172]. The single amino acid substitution S191T found in the extracellular region of the NK1.1 antigen is responsible for the strain-dependent alloreactivity of the PK136 antibody, as demonstrated by sequence alignment and epitope mapping recently[173].

The identification of ligands for NKR-P1 receptors has been a subject of controversy. Early investigation into the binding of rat NKR-P1A showed high binding affinity of the protein to several monosaccharides such as GlcNAc and GalNAc, and the binding activity was reported to be Ca^{2+} dependent[174]. Oligosaccharides derived from heparin, oligosaccharide sequences of the blood group family, the ganglio family, and glycosaminglycans were further described as new high affinity ligands for NKR-P1A with significant potential in antitumor therapy[175]. Although these findings have never been reproduced by other groups, and two additional re-evaluations of carbohydrate binding activity for NKR-P1A were published[176,177], the research in the field of saccharide ligands for NKR-P1A continued and number of compounds such as derivates of aminosugars, chitooligomers, highly branched glycoclusters, and peptides were reported to be prominent classes of high affinity ligands for NKR-P1A[178-188]. However, a recent announcement from the Ethical Committee of the Institute of Microbiology and Charles University in Prague indicates that these findings were probably fabricated.

Protein ligands for two members of mouse NKR-P1 family have been identified. While the activating mouse NKR-P1F has been shown to bind Clr-g protein[43], the inhibitory mouse NKR-P1B/D has been reported to recognize Clr-b protein[189]. Surprisingly, both Clr-g and Clr-b are not MHC class I related molecules, but instead are members of the Clr family, which represent type II transmembrane C-type lectin-like proteins[190]. The cDNA of *Clr* was previously known from osteoclastogenesis, therefore the members of the gene family were also referred as osteoclast inhibitory lectins (Ocil)[191,192]. From a genetic point of view, the *NKR-P1* genes are remarkably intertwined in the mouse NKC with the *Clr* genes encoding their ligands[43]. Investigation of expression level of Clr-b revealed its dramatic down-regulation on tumor cell lines, which is consistent with the missing-self hypothesis. Moreover, these findings broaden the missing-self hypothesis because in this case, the control of NK cell activation is MHC I independent[43,189].

In humans, NKR-P1A receptor has been demonstrated to bind LLT1, which represents the human ortholog of mouse Clr[193,194]. Whereas LLT1 was observed to enhance IFN-γ production by human T cells and NK cells, NKR-P1A have been demonstrated either to induce or inhibit NK cell mediated cytotoxicity. However, these variable results require future attention[195-197]. From a structural point of view, no three dimensional structure of any of the NKR-P1 family receptors has been published so far.

CHAPTER 2
AIMS OF THE THESIS

The primary goal of this thesis is to contribute to ongoing research of NKR-P1 receptors and fill the gaps in knowledge on their structure as a key for understanding their function. The aims of thesis include:

- To develop methods for expression, refolding, and purification of mouse NKR-P1 receptors providing sufficient amount of proteins for structural studies
- To construct refined homology models for mouse NKR-P1 receptors
- To study the structure of mouse NKR-P1 receptors by protein crystallography
- To study the solution structure of mouse NKR-P1 receptors by mass spectrometric techniques

CHAPTER 3
METHODS

The papers in this Ph.D. thesis include a detailed description of the methods, including all the technical details necessary for their reproducibility. Therefore, Chapter 3 presents only list of research methods used in this thesis.

List of research methods:

- Gene cloning and DNA manipulation
- Protein expression, refolding *in vitro,* and protein purification
- Disulfide bond mapping
- Analytical ultracentrifugation
- Dynamic light scattering
- Differential scanning calorimetry
- Raman and infrared spectroscopy
- Chemical cross-linking coupled to mass spectrometry
- Hydrogen/Deuterium exchange coupled to mass spectrometry
- Native mass spectrometry and ion mobility measurement
- Sequence and phylogeny analysis
- Protein homology modeling
- Protein crystallization and protein crystallography
- Protein NMR spectroscopy

CHAPTER 4
RESULTS AND DISCUSSION

The main focus of this thesis is to contribute to ongoing research of NKR-P1 receptors and fill the gaps in knowledge on their structure as a key for understanding their function. Although the NKR-P1 receptors were among the first surface NK receptors identified on rodent NK cells more than 20 years ago[198,199], there is still very little known about their biological function as well as their physiological ligand. Furthermore, no three-dimensional structure of any of the NKR-P1 family has been reported so far. Nevertheless, in recent years, there has been a renewal of interest in these receptors due to their unique role in the MHC class I independent missing-self recognition[189]. This extensive immunological research has shed some light on the function of the several NKR-P1 receptors; however, recombinant expression providing sufficient amount for structural characterization still remained an enigma.

This fact prompted us to develop a method for recombinant expression of soluble NKR-P1 receptors. Since the previous immunological research was related to the mouse NKR-P1 family, our effort to express recombinant proteins was directed towards those orthologs. In the paper *Rozbesky et al*[200], we described a simple and efficient protocol for the production of the soluble C-type lectin-like domain of mouse NKR-P1C. The single C-type lectin-like domain of NKR-P1C was expressed because the desired biological activity responsible for ligand binding seemed to be supported by this discrete domain. Furthermore, at least two transcriptional isoforms of NKR-P1C have recently been described with differences in the sequence of the stalk region, while their C-type lectin-like domains are the same[201]. In contrast to the protocols reported previously for rat NKR-P1A[174,202], we decided to design the expression construct without any targeting signals or purification tags.

In the first step, cDNA was obtained by reverse transcription of total RNA isolated from mouse spleen. The gene encoding NKR-P1C was amplified by PCR and finally subcloned into a pET-30a(+) expression vector. Upon induction of bacterial expression by IPTG, the protein precipitated into inclusion bodies from which it could be refolded *in vitro* by a rapid dilution method. The protein refolding was the most critical step in protein preparation and required significant optimization. The key parameters for

refolding were pH 8.5 and a high concentration of arginine, which in general decreases aggregation by slowing the rate of protein-protein interactions. Another key parameter was a cysteamine/cystamine redox shuffling pair required for correct disulfide reoxidation. After refolding, the protein was dialyzed and purified by one step anion exchange chromatography on Q Sepharose.

The protein identity was verified by high resolution ESI FT ICR mass spectrometry. The monoisotopic mass of intact protein (**Fig. 8a**) was in complete agreement with mass of expected amino acid sequence with three intramolecular disulfide bonds. Furthermore, mass spectrometry demonstrated a high purity of protein sample that reached more than 98%. Native size of soluble NKR-P1C was examined by size exclusion chromatography and analytical ultracentrifugation. Chromatography on the Superdex 75 column suggested that the protein eluted exclusively as a monomer. This finding was further corroborated by sedimentation velocity as well as sedimentation equilibrium experiments. In comparison, protein samples of homologous proteins such as human CD69 were observed to be noncovalent dimers[203]. Since the ligand of NKR-P1C remains elusive, we were not able to investigate its biological activity. Instead, the protein folding was analyzed by NMR spectroscopy. A ^{1}H-^{15}N HSQC spectrum of ^{15}N labeled protein (**Fig. 8b**) demonstrated a good dispersion of the backbone and side chain signals with excellent 5 ppm chemical shift dispersion in proton dimension, indicating a compact fold of NKR-P1C.

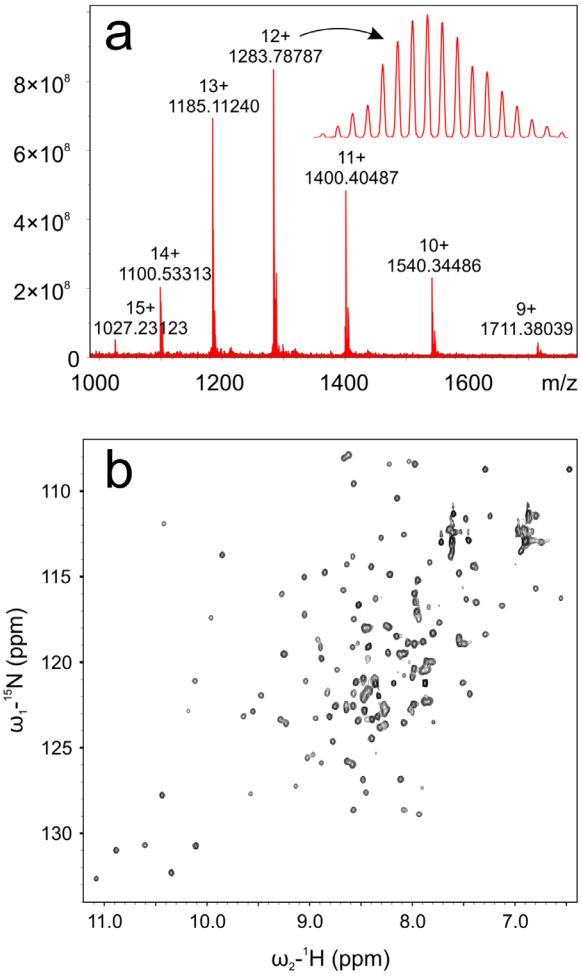


Figure 8 *Mass spectrum of multiply charged recombinant mouse NKR-P1C (**a**), and ^{1}H-^{15}N HSQC spectrum of the uniformly ^{15}N labaled NKR-P1C indicating a good dispersion of backbone and side chain-signals (**b**)[200].*

Since the disulfide arrangement is extremely conserved in C-type lectin-like domains and it is essential for their biological activity, we have mapped the disulfide bonds in a protein sample of NKR-P1C in order to further investigate the fold quality. The disulfide bond arrangement in NKR-P1C was determined by mass spectrometry according to the previously published protocol[204]. Briefly, the protein was separated by SDS-PAGE and digested by trypsin or Asp-N protease. The peptide mixtures were desalted and analyzed by high resolution LC-MS and the disulfide bonds were finally identified by Links software. The identified disulfide arrangement in NKR-P1C was in agreement with the highly conserved disulfide pattern of C-type lectin-like domains[114] which was further evidence of correct NKR-P1C folding.

In contrast to the protocols reported previously for rat NKR-P1A[174,202], our optimized method for preparation of mouse NKR-P1C provided a high yield: approximately 20 mg of pure protein per liter of bacterial culture. In addition to the protocol, the recombinant protein can be easily prepared over a short period of 4-6 days using common biochemical laboratory equipment.

In the paper *Sovová et al*[205], we generated refined homology models for C-type lectin-like domains of mouse NKR-P1A, NKR-P1C, NKR-P1F, and NKR-P1G. In the first step, candidate template sequences were identified by BLAST and the template structures with high sequence identity were used for modeling using Modeller 9v4. For further analysis, the best model out of 10 calculated for every receptor, was chosen by the distribution of residues in the Ramachandran plot and the stereochemical g-factor calculated by Procheck. Finally, all models were refined by at least 10 ns of molecular dynamic simulations.

The overall fold of all models (**Fig. 9**) was very similar to the C-type lectin-like domain containing α-helices surrounding a ß core composed of antiparalell ß-strands. The models contained three disulfide bonds in the arrangement described previously[200]. Although the core of the protein was almost the same for all models, the topological organization of secondary elements slightly varied among the models. In addition to this, topology differences were observed for prominent loops mainly anchored by small antiparalell sheets. These differences in the spatial arrangement of the loop are also reflected in the behavior of each model in molecular dynamics simulations. While the

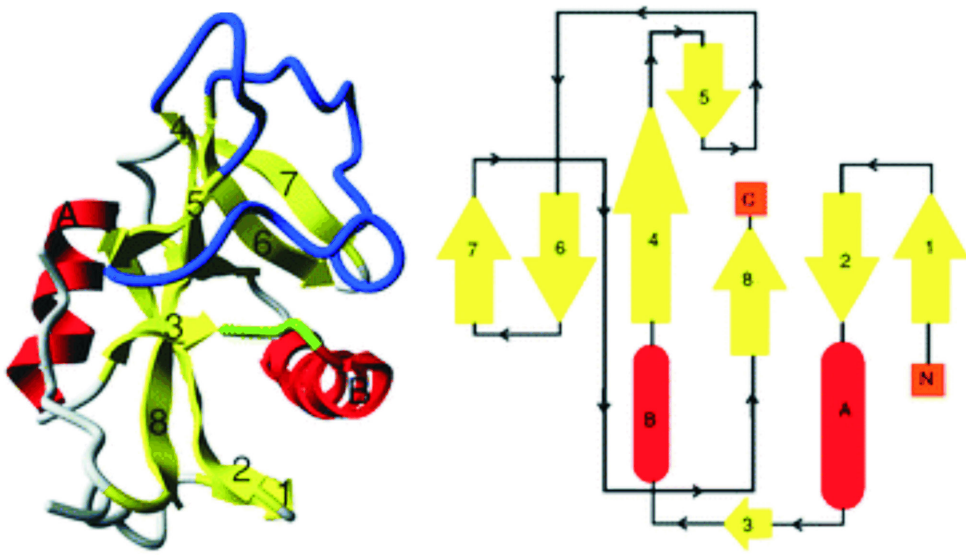


Figure 9 *Representative homology model of mouse NKR-P1C with the fold similar to the C-type lectin-like domain and the conserved loop region in blue color*[205].

protein core was stable in all simulations with C_α rmsds reaching plateau values, the root mean square fluctuations for C_α calculated from the last nanosecond of the equilibrated part of each trajectory clearly identified the long loop region as very flexible. The root mean square fluctuations of the residues involved in the loops were higher than the average for the whole protein, although the structure within the loop was stable in a time period that could be simulated. The loop occasionally folded back onto the protein surface, where it formed a hydrophobic interaction with the two tryptophan residues which are highly conserved in the WKW sequence motif. In conclusion, these results suggested that the loop region is stably anchored to the core, however, it can adopt alternative positions.

It is noteworthy that the sequence analysis of the 33 available C-type lectin-like domains among NKR-P1 orthologs revealed that the region with lowest conservation corresponds to the long loop region as described previously. Substitutions in the loop (L160-T184) cause significant changes in chemical character of some loop residues among natural NKR-P1 variants. It leads to two groups typified by NKR-P1subfamilies: A-D and F/G. The NKR-P1A-D group presents particularly L160, N164, T172, K179 and T184, while the NKR-P1F/G group substitutes a polar residue (E or Q) for L160, a ß -branched residue (T, V or I) for N164, a more hydrophobic residue (I or V) for T172, Ser

or Arg for K178, and a more polar residue (N, D or E) for T184. The *in vivo* binding experiments with Clr ligands[161] suggest that these sequence differences in the long loop may be related to ligand specificity.

The C-type lectin-like domain of mouse NKR-P1A was further prepared according to the protocol for mouse NKR-P1C described previously[200] and both NKR-P1A and NKR-P1C were structurally analyzed by Raman and FTIR spectroscopy. The constructed models were in good agreement with the Raman and infrared spectra not only in terms of overall secondary structure content, but the models also captured key structural and dynamics features consistent with dynamics simulations.

Both mouse NKR-P1A and NKR-P1C were crystallized using the hanging drop vapor diffusion method. While crystallization of NKR-P1C was unsuccessful, NKR-P1A provided tetragonal monocrystals suitable for diffraction. Diffraction data were collected from a single crystal at synchrotron radiation source Bessy II in Helmholtz Zentrum Berlin. The phase problem was solved by molecular replacement using the extracellular domain of CD69 as a search model and the structure was refined to 1.7 Å resolution and very good agreement with the experimental data. In addition to this, the determined structure of NKR-P1A (**Fig. 10a**) described in *Kolenko & Rozbesky et al*[206] is the first known structure in the NKR-P1 family.

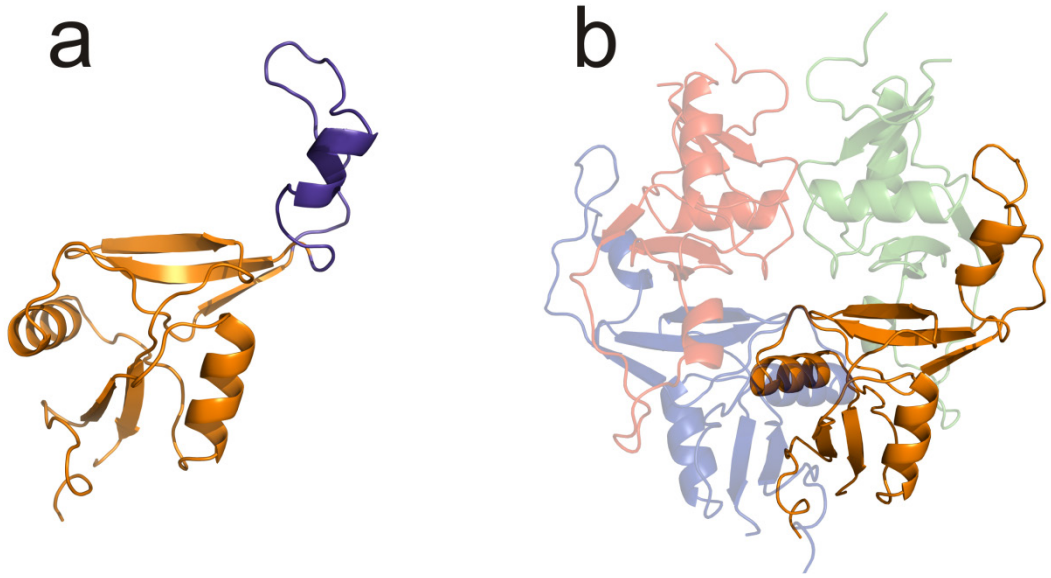


Figure 10 *Crystal structure of one chain of mouse NKR-P1A with extended long loop region in purple color (**a**). Two types of dimer observed in the crystal are both of interest: the first one utilizing the swapped loops, and the second one reflecting the human CD69 dimer (**b**).*

The overall fold of NKR-P1A was similar to that of CD69 but differed significantly in the long loop region (Y158-D187) which was surprisingly extended from the compact core. The compact core consisted of two α-helices and two antiparallel ß-sheets and was topologically consistent with the classical C-type lectin-like domain[114]. Furthermore, arrangement of disulfide bonds was in agreement with the conserved disulfide pattern of C-type lectin-likes proteins. It is noteworthy that the extended loop pointing away from the globular domain exhibited domain swapping with a symmetry related molecule. This extended loop has already been observed in other C-type lectin-like proteins but not in any NK receptors. A well documented example of the extended loop is represented by the macrophage mannose receptor, however, its biological implications have not been evaluated[207]. On the other hand, bitiscetin, a C-type lectin snake venom inducing platelet agglutination, binds a von Willebrand factor A1 domain into the central groove of its domain-swapped dimer which is similar to the dimer formed in the NKR-P1A crystal structure[208]. Indeed, numerous C-type lectin-like proteins of snake venom modulating the homeostatic system and forming domain-swapped dimers and oligomers have been reported so far[209-211]. In the same year as we published the structure of NKR-P1A, the crystal structure of the human C-type lectin-like related macrophage receptor CLEC5A was reported by *Watson et al*[212]. Although the crystal structure of this receptor, which binds to the dengue virus, was consistent with the classical C-type lectin-like domain[114] and its loop was associated with the protein core, molecular dynamic simulations suggested a significant conformational flexibility of the loop which could participate in ligand binding. Remarkably, these findings are very similar to the molecular dynamic simulation and structural aspects of NKR-P1A as we described previously[205,206].

In addition to the NKR-P1A, two types of dimers (**Fig. 10b**) were observed in the crystal structure: the stronger type A utilizing the swapped loops, and type B reflecting the human CD69. While the A dimer had its N-termini placed on the opposite ends, the B dimer had the termini located on the one side of the dimer. The interaction in the A dimer corresponded to the domain swapping effect of the extended loops with total interface area of about 1700 $Å^2$ involving hydrogen bonds and hydrophobic interactions. Although these findings suggest the potential participation of the loop in interactions with ligands, further investigation is necessary to confirm this hypothesis.

In the paper *Kolenko et al*[213], we identified a novel point mutation in the gene for mouse NKR-P1A leading to the substitution of His at position 107 by Arg. The recombinant H107R variant was expressed, purified and crystallized as the wild type. The crystal structure of H107R NKR-P1A has been determined by X-ray diffraction at 2.3 Å resolution from merohedrally twinned crystals. The structure contained two protein chains with three phosphate ions attached at the interface of the molecules. The quality of the observed electron density did not allow the localization of all residues, particularly Y149, P161 and D162 in chain B. Whereas two phosphate ions were bound close to the interface between the monomers in asymmetric unit, the third phosphate ion was localized between R177 belonging to opposite chains. The rmsd between C_α atoms of the wild type and chains A and B of the H107R mutant were 0.6 and 1.8 Å, respectively. Thus the overall structure of the mutant did not differ significantly from the wild type structure. Nevertheless, the mutation and subsequent binding of phosphate ions at the interface probably played a role in the twinning of the crystals in the H107R mutant.

The crystal structure of NKR-P1A evoked several questions about the flexibility and topology of the loop region in solution. Therefore, in the paper *Rozbesky et al*[214], mass spectrometric techniques such as chemical cross-linking and H/D exchange to investigate the NKR-P1A structure in solution were applied.

In chemical cross-linking, generally, the residues which are in close proximity to each other may be converted into covalent bonds by a cross-linker. The cross-linked residues which are then efficiently identified by mass spectrometry provide experimental distance constraints which may be applied for modeling purposes. In order to study the loop conformation in solution, NKR-P1A was cross-linked with the carboxyl-amine coupling reagent EDC and two homobifunctional amine-amine coupling reagents DSS and DSG. In the first step, optimization of reaction conditions was performed to get the highest relative yield of protein containing only a single intramolecular cross-link, which minimized the likelihood of distorting the tertiary structure. The cross-linked NKR-P1A was then separated by SDS PAGE and digested *in gel* by different endoproteases. Generated peptides were desalted and loaded onto a reversed phase column coupled online to an ESI FT ICR mass spectrometer. The high mass accuracy was sufficient to provide unambiguous assignment of cross-links using the Links software.

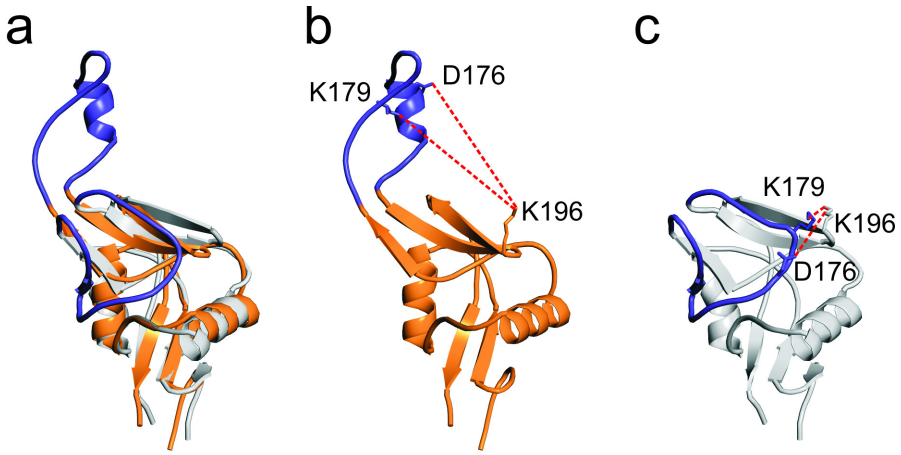


Figure 11 *Cross-linking of NKR-P1A combined with mass spectrometry. (**a**) The different conformations of the loop region (purple) in a structural alignment of the crystal structure (orange) and the homology model (gray) of NKR-P1A. (**b**) The spatial distances (red dashed lines) between K179-K196 and D176-K196 in the crystal structure of NKR-P1A are inconsistent with cross-linker lengths. (**c**) The spatial distance constraints for K179-K196 and D176-K196 obtained in this study are consistent with the inter-residue distances in the homology model*[214].

The inter-residue distances obtained in this study were in excellent agreement with the inter-residue distances in the crystal structure with exception of K179-K196 and D176-K196 (**Fig. 11**). Based on the length of spacer arm in the DSG, the C_α-C_α distance between DSG cross-linked K179 from the loop region and K196 from the compact core of NKR-P1A should be up to 20 Å, but in the crystal structure it was 29.6 Å. Furthermore, the C_α-C_α distance of 32.9 Å between D176 from the loop region and K196 from the core in the crystal structure was also inconsistent with the ability of EDC to cross-link these residues. On the contrary, all experimental distance constraints fitted the distances in the homology model of NKR-P1A described previously. These findings showed that in solution, the loop region of NKR-P1A is not extended from the compact core, as observed in the crystal structure, but is in close proximity to the core as suggested by the homology model. Further, the absence of cross-links between K179 or D176 and lysines or acidic residues, which are spread throughout the core of NKR-P1A, suggests that the backbone in the loop region is not particularly flexible.

To investigate the NKR-P1A loop conformation by other approach, H/D exchange coupled to mass spectrometry was applied. In general, H/D exchange provides

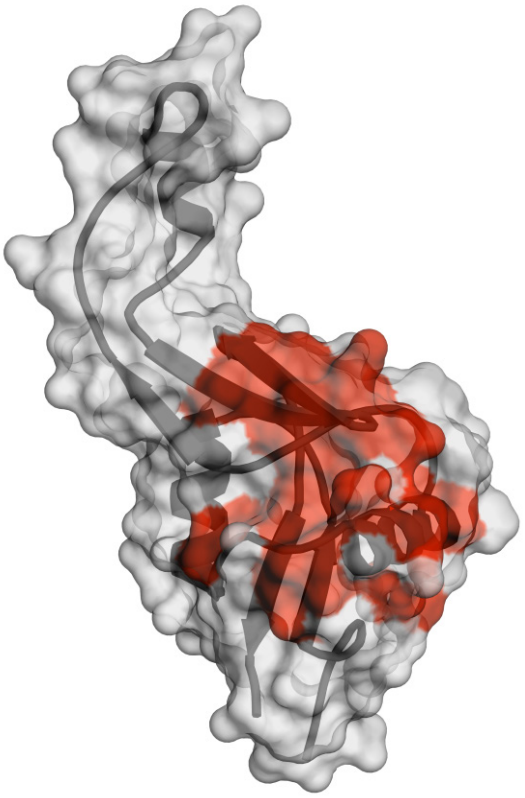


Figure 12 *Differences in amide hydrogen exchange rates between NKR-P1A and NKR-P1A^{NL} mapped onto the crystal structure of NKR-P1A. The regions with a decreased exchange rate in NKR-P1A are in red*[214].

information on the local solvent accessibility of proteins that is valuable in studying protein dynamics, structure, and function. The aim of this analysis was to compare the kinetics of H/D exchange for NKR-P1A and NKR-P1A in which the loop was removed and replaced with two alanines (NKR-P1A^{NL}). For NKR-P1A^{NL} expression, the construct was prepared by PCR mediated deletion–insertion mutagenesis. NKR-P1A and NKR-P1A^{NL} were then subjected to H/D exchange followed by reduction of disulfide bonds with TCEP. Reduced samples were loaded onto a semi-automated chromatographic system for pepsin digestion, peptide desalting, and peptide separation on a reversed phase column. The extent of deuterium incorporation into each peptide was analyzed by high resolution ESI FT ICR mass spectrometry. When compared to NKR-P1A^{NL}, NKR-P1A showed slower deuteration rates for regions 130–140, 138–143, and 191–199. This analysis suggested that these regions are protected by the loop (**Fig. 12**), which is consistent with the cross-linking study and the homology model. In addition to this, our findings were further corroborated by NMR spectroscopy.

Taken together, cross-linking in combination with H/D exchange as well as NMR spectroscopy revealed that the solution structure differs from the crystal structure in the conformation of the conserved loop. While the conserved loop was in close proximity to compact core in solution, it was extended from the core in the crystal structure where it interacted with the surface of a symmetry-related molecule. This suggests a testable hypothesis that the conformational change of this loop may be induced by ligand binding.

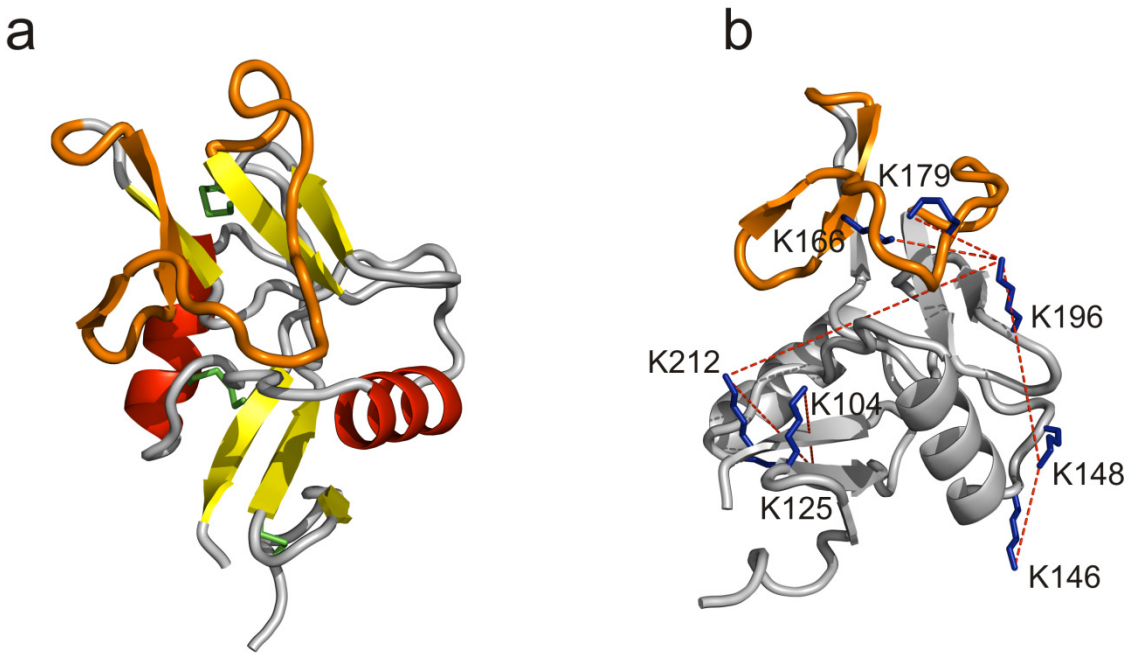


Figure 13 *Structural model of NKR-P1C with the loop region colored in orange and disulfide bonds colored in green (**a**). Model of NKR-P1C with lysine residues involved in chemical cross-linking. Cross-linked residues are highlighted in dashed line (**b**)*[215].

Since the crystallization of NKR-P1C failed in the paper *Rozbesky et al*[215], we decided to use set of methods, ranging from chemical cross-linking with sequence threading, restraint-based molecular modeling and steered molecular dynamics, to ion mobility mass spectrometry in order to describe the structure of NKR-P1C in solution (**Fig. 13a**).

In chemical cross-linking experiments, both cross-linkers were applied as a 1:1 mixture of nondeuterated and deuterated (d_0/d_4) derivate, which finally resulted in characteristic isotopic pattern of peptides bearing the cross-linker. This specific signature, together with high mass accuracy greatly facilitated unambiguous assignment of cross-links. Using this approach, 7 cross-links formed by DSS and 3 cross-links formed by DSG were identified. Within the whole cross-linking network (**Fig. 13b**) only K166 and K179 were a part of the region related to loop region in homologous NKR-P1A, whereas the other lysine residues were located in the globular domain. The experimental constraints matched the conformation in which the loop was packed onto the protein surface, similarly to the homology model of NKR-P1C described previously[205] or to the dengue virus receptor CLEC5A[212].

Therefore the homology modeling was performed using the recently published crystal structure of mouse NKR-P1A as a template for the protein core and the published

structure of human CLEC5A for the loop region. Constraints for the loop region coming from the experimentally determined distances were solved by a non-equilibrium molecular dynamics simulation using YASARA, applying an external force on the loop atoms in the direction of coordinates within the experimentally determined range. Confrontation of final model with cross-linking data showed that all experimental distances were in agreement with the model.

In order to verify the model of NKR-P1C, native mass spectrometry with ion mobility measurements was performed. The arrival time distributions strongly suggested that the protein was present in a unique, very homogenous conformation. The calculated value of collisional cross section (CCS) of 1444 $Å^2$ was very close to the theoretical CCS of 1411 $Å^2$ for our model. In comparison, the homologue NKR-P1A with loop extended showed a much higher theoretical CCS of 1618 $Å^2$.

In conclusion, we showed that the structure of NKR-P1C in solution is similar to the homologous protein NKR-P1A and that its loop region is attached to the protein core. In addition to this, this was the first report describing combined use of chemical cross-linking, homology modeling, and ion mobility mass spectrometry for protein structure determination.

In conclusion of this thesis, we have developed a novel method for expression, refolding, and purification of mouse C-type lectin-like domain of NKR-P1 receptors. Homology modeling, sequence, and phylogenetic analysis suggested the key role of the conserved loop in the structure of NKR-P1 receptors that might contribute to ligand binding specificity and consequently to differentiation between activating and inhibitory receptors. This hypothesis was partially corroborated by discrepancy between the structure of NKR-P1A in crystal and solution which indicated its concrete dynamical behavior in ligand binding. In addition to this, the crystal structure of NKR-P1A is the first known structure in the NKR-P1 family. Further investigations are necessary to clarify the molecular details of these interactions and their relevance to those occurring at the cellular level. Finally, from the point of protein structure determination, we implemented a novel combination of mass spectrometric techniques such as chemical cross-linking, H/D exchange, and ion mobility measurements to study protein structures of NKR-P1A and NKR-P1C in solution.

CHAPTER 5
SUMMARY

- A new method for expression, refolding, and purification of mouse NKR-P1 receptors has been developed.
- Refined homology models have been constructed for NKR-P1 receptors.
- The crystal structure of mouse NKR-P1A has been determined by molecular replacement to 1.7Å resolution.
- In the crystal structure of mouse NKR-P1A, the overall fold was similar to the other C-type lectin-like domains but differed significantly in the unique conformation of the loop which was extended from the compact core.
- The crystal structure of the H107R variant of mouse NKR-P1A has been determined to 2.3 Å resolution.
- The mutation H107R does not have a significant impact on the overall structure of mouse NKR-P1A.
- The solution structure of mouse NKR-P1A has been determined using the combination of chemical cross-linking and H/D exchange coupled to high resolution mass spectrometry.
- The loop extended from the compact core of NKR-P1A in the crystal structure was found to be closely attached to the protein core in solution.
- Hypothesis for the loop participation in ligand binding has been suggested based on the discrepancy between crystal and solution structure.
- The solution structural model of mouse NKR-P1C has been determined using the combination of chemical cross-linking, homology modeling, and ion mobility measurement.

REFERENCES

1. Kiessling, R., Klein, E. & Wigzell, H. "Natural" killer cells in the mouse. I. Cytotoxic cells with specificity for mouse Moloney leukemia cells. Specificity and distribution according to genotype. *Eur J Immunol* **5**, 112-7 (1975).
2. Sun, J.C., Beilke, J.N. & Lanier, L.L. Adaptive immune features of natural killer cells. *Nature* **457**, 557-61 (2009).
3. Sun, J.C. & Lanier, L.L. Natural killer cells remember: an evolutionary bridge between innate and adaptive immunity? *Eur J Immunol* **39**, 2059-64 (2009).
4. Strowig, T., Brilot, F. & Munz, C. Noncytotoxic functions of NK cells: direct pathogen restriction and assistance to adaptive immunity. *J Immunol* **180**, 7785-91 (2008).
5. Yokoyama, W.M. Mistaken notions about natural killer cells. *Nat Immunol* **9**, 481-5 (2008).
6. Smyth, M.J., Swann, J., Cretney, E., Zerafa, N., Yokoyama, W.M. & Hayakawa, Y. NKG2D function protects the host from tumor initiation. *J Exp Med* **202**, 583-8 (2005).
7. Imai, K., Matsuyama, S., Miyake, S., Suga, K. & Nakachi, K. Natural cytotoxic activity of peripheral-blood lymphocytes and cancer incidence: an 11-year follow-up study of a general population. *Lancet* **356**, 1795-9 (2000).
8. Coca, S., Perez-Piqueras, J., Martinez, D., Colmenarejo, A., Saez, M.A., Vallejo, C., Martos, J.A. & Moreno, M. The prognostic significance of intratumoral natural killer cells in patients with colorectal carcinoma. *Cancer* **79**, 2320-8 (1997).
9. Eckl, J., Buchner, A., Prinz, P.U., Riesenberg, R., Siegert, S.I., Kammerer, R., Nelson, P.J. & Noessner, E. Transcript signature predicts tissue NK cell content and defines renal cell carcinoma subgroups independent of TNM staging. *J Mol Med (Berl)* **90**, 55-66 (2012).
10. Carrega, P., Morandi, B., Costa, R., Frumento, G., Forte, G., Altavilla, G., Ratto, G.B., Mingari, M.C., Moretta, L. & Ferlazzo, G. Natural killer cells infiltrating human nonsmall-cell lung cancer are enriched in CD56 bright CD16(-) cells and display an impaired capability to kill tumor cells. *Cancer* **112**, 863-75 (2008).
11. Platonova, S., Cherfils-Vicini, J., Damotte, D., Crozet, L., Vieillard, V., Validire, P., Andre, P., Dieu-Nosjean, M.C., Alifano, M., Regnard, J.F., Fridman, W.H., Sautes-Fridman, C. & Cremer, I. Profound coordinated alterations of intratumoral NK cell phenotype and function in lung carcinoma. *Cancer Res* **71**, 5412-22 (2011).
12. Diefenbach, A., Jensen, E.R., Jamieson, A.M. & Raulet, D.H. Rae1 and H60 ligands of the NKG2D receptor stimulate tumour immunity. *Nature* **413**, 165-71 (2001).
13. Sivori, S., Parolini, S., Marcenaro, E., Castriconi, R., Pende, D., Millo, R. & Moretta, A. Involvement of natural cytotoxicity receptors in human natural killer cell-mediated lysis of neuroblastoma and glioblastoma cell lines. *J Neuroimmunol* **107**, 220-5 (2000).
14. Fauriat, C., Just-Landi, S., Mallet, F., Arnoulet, C., Sainty, D., Olive, D. & Costello, R.T. Deficient expression of NCR in NK cells from acute myeloid leukemia: Evolution during leukemia treatment and impact of leukemia cells in NCRdull phenotype induction. *Blood* **109**, 323-30 (2007).
15. Koh, C.Y., Blazar, B.R., George, T., Welniak, L.A., Capitini, C.M., Raziuddin, A., Murphy, W.J. & Bennett, M. Augmentation of antitumor effects by NK cell inhibitory receptor blockade in vitro and in vivo. *Blood* **97**, 3132-7 (2001).
16. Shahied, L.S., Tang, Y., Alpaugh, R.K., Somer, R., Greenspon, D. & Weiner, L.M. Bispecific minibodies targeting HER2/neu and CD16 exhibit improved tumor lysis when placed in a divalent tumor antigen binding format. *J Biol Chem* **279**, 53907-14 (2004).
17. Ruggeri, L., Capanni, M., Urbani, E., Perruccio, K., Shlomchik, W.D., Tosti, A., Posati, S., Rogaia, D., Frassoni, F., Aversa, F., Martelli, M.F. & Velardi, A. Effectiveness of donor natural killer cell alloreactivity in mismatched hematopoietic transplants. *Science* **295**, 2097-100 (2002).
18. Giebel, S., Locatelli, F., Lamparelli, T., Velardi, A., Davies, S., Frumento, G., Maccario, R., Bonetti, F., Wojnar, J., Martinetti, M., Frassoni, F., Giorgiani, G., Bacigalupo, A. & Holowiecki, J. Survival advantage with KIR ligand incompatibility in hematopoietic stem cell transplantation from unrelated donors. *Blood* **102**, 814-9 (2003).
19. Miller, J.S., Cooley, S., Parham, P., Farag, S.S., Verneris, M.R., McQueen, K.L., Guethlein, L.A., Trachtenberg, E.A., Haagenson, M., Horowitz, M.M., Klein, J.P. & Weisdorf, D.J. Missing KIR ligands are associated with less relapse and increased graft-versus-host disease (GVHD) following unrelated donor allogeneic HCT. *Blood* **109**, 5058-61 (2007).
20. Davies, S.M., Ruggieri, L., DeFor, T., Wagner, J.E., Weisdorf, D.J., Miller, J.S., Velardi, A. & Blazar, B.R. Evaluation of KIR ligand incompatibility in mismatched unrelated donor hematopoietic

transplants. Killer immunoglobulin-like receptor. *Blood* **100**, 3825-7 (2002).
21. Lee, S.H., Miyagi, T. & Biron, C.A. Keeping NK cells in highly regulated antiviral warfare. *Trends Immunol* **28**, 252-9 (2007).
22. Scalzo, A.A., Corbett, A.J., Rawlinson, W.D., Scott, G.M. & Degli-Esposti, M.A. The interplay between host and viral factors in shaping the outcome of cytomegalovirus infection. *Immunol Cell Biol* **85**, 46-54 (2007).
23. Loh, J., Chu, D.T., O'Guin, A.K., Yokoyama, W.M. & Virgin, H.W.t. Natural killer cells utilize both perforin and gamma interferon to regulate murine cytomegalovirus infection in the spleen and liver. *J Virol* **79**, 661-7 (2005).
24. Warfield, K.L., Perkins, J.G., Swenson, D.L., Deal, E.M., Bosio, C.M., Aman, M.J., Yokoyama, W.M., Young, H.A. & Bavari, S. Role of natural killer cells in innate protection against lethal ebola virus infection. *J Exp Med* **200**, 169-79 (2004).
25. Smith, H.R., Heusel, J.W., Mehta, I.K., Kim, S., Dorner, B.G., Naidenko, O.V., Iizuka, K., Furukawa, H., Beckman, D.L., Pingel, J.T., Scalzo, A.A., Fremont, D.H. & Yokoyama, W.M. Recognition of a virus-encoded ligand by a natural killer cell activation receptor. *Proc Natl Acad Sci U S A* **99**, 8826-31 (2002).
26. Arase, H., Mocarski, E.S., Campbell, A.E., Hill, A.B. & Lanier, L.L. Direct recognition of cytomegalovirus by activating and inhibitory NK cell receptors. *Science* **296**, 1323-6 (2002).
27. Vidal, S.M. & Lanier, L.L. NK cell recognition of mouse cytomegalovirus-infected cells. *Curr Top Microbiol Immunol* **298**, 183-206 (2006).
28. Voigt, S., Mesci, A., Ettinger, J., Fine, J.H., Chen, P., Chou, W. & Carlyle, J.R. Cytomegalovirus evasion of innate immunity by subversion of the NKR-P1B:Clr-b missing-self axis. *Immunity* **26**, 617-27 (2007).
29. Piccioli, D., Sbrana, S., Melandri, E. & Valiante, N.M. Contact-dependent stimulation and inhibition of dendritic cells by natural killer cells. *J Exp Med* **195**, 335-41 (2002).
30. Morandi, B., Bougras, G., Muller, W.A., Ferlazzo, G. & Munz, C. NK cells of human secondary lymphoid tissues enhance T cell polarization via IFN-gamma secretion. *Eur J Immunol* **36**, 2394-400 (2006).
31. Martin-Fontecha, A., Thomsen, L.L., Brett, S., Gerard, C., Lipp, M., Lanzavecchia, A. & Sallusto, F. Induced recruitment of NK cells to lymph nodes provides IFN-gamma for T(H)1 priming. *Nat Immunol* **5**, 1260-5 (2004).
32. Takeda, K. & Dennert, G. The development of autoimmunity in C57BL/6 lpr mice correlates with the disappearance of natural killer type 1-positive cells: evidence for their suppressive action on bone marrow stem cell proliferation, B cell immunoglobulin secretion, and autoimmune symptoms. *J Exp Med* **177**, 155-64 (1993).
33. Bolovan-Fritts, C.A. & Spector, S.A. Endothelial damage from cytomegalovirus-specific host immune response can be prevented by targeted disruption of fractalkine-CX3CR1 interaction. *Blood* **111**, 175-82 (2008).
34. Yoneda, O., Imai, T., Goda, S., Inoue, H., Yamauchi, A., Okazaki, T., Imai, H., Yoshie, O., Bloom, E.T., Domae, N. & Umehara, H. Fractalkine-mediated endothelial cell injury by NK cells. *J Immunol* **164**, 4055-62 (2000).
35. Croy, B.A., van den Heuvel, M.J., Borzychowski, A.M. & Tayade, C. Uterine natural killer cells: a specialized differentiation regulated by ovarian hormones. *Immunol Rev* **214**, 161-85 (2006).
36. Hanna, J., Goldman-Wohl, D., Hamani, Y., Avraham, I., Greenfield, C., Natanson-Yaron, S., Prus, D., Cohen-Daniel, L., Arnon, T.I., Manaster, I., Gazit, R., Yutkin, V., Benharroch, D., Porgador, A., Keshet, E., Yagel, S. & Mandelboim, O. Decidual NK cells regulate key developmental processes at the human fetal-maternal interface. *Nat Med* **12**, 1065-74 (2006).
37. Vivier, E., Tomasello, E., Baratin, M., Walzer, T. & Ugolini, S. Functions of natural killer cells. *Nat Immunol* **9**, 503-10 (2008).
38. Karre, K., Ljunggren, H.G., Piontek, G. & Kiessling, R. Selective rejection of H-2-deficient lymphoma variants suggests alternative immune defence strategy. *Nature* **319**, 675-8 (1986).
39. Ljunggren, H.G. & Karre, K. In search of the 'missing self': MHC molecules and NK cell recognition. *Immunol Today* **11**, 237-44 (1990).
40. Bryceson, Y.T., March, M.E., Ljunggren, H.G. & Long, E.O. Synergy among receptors on resting NK cells for the activation of natural cytotoxicity and cytokine secretion. *Blood* **107**, 159-66 (2006).
41. Oberg, L., Johansson, S., Michaelsson, J., Tomasello, E., Vivier, E., Karre, K. & Hoglund, P. Loss or mismatch of MHC class I is sufficient to trigger NK cell-mediated rejection of resting lymphocytes in vivo - role of KARAP/DAP12-dependent and -independent pathways. *Eur J Immunol* **34**, 1646-53 (2004).
42. Lanier, L.L. NK cell recognition. *Annu Rev Immunol* **23**, 225-74 (2005).
43. Iizuka, K., Naidenko, O.V., Plougastel, B.F., Fremont, D.H. & Yokoyama, W.M. Genetically linked C-type lectin-related ligands for the NKRP1 family of natural killer cell receptors. *Nat Immunol* **4**, 801-7 (2003).
44. Bauer, S., Groh, V., Wu, J., Steinle, A., Phillips, J.H., Lanier, L.L. & Spies, T. Activation of NK cells

and T cells by NKG2D, a receptor for stress-inducible MICA. *Science* **285**, 727-9 (1999).
45. Raulet, D.H. & Guerra, N. Oncogenic stress sensed by the immune system: role of natural killer cell receptors. *Nat Rev Immunol* **9**, 568-80 (2009).
46. Raulet, D.H. Roles of the NKG2D immunoreceptor and its ligands. *Nat Rev Immunol* **3**, 781-90 (2003).
47. Brandt, C.S., Baratin, M., Yi, E.C., Kennedy, J., Gao, Z., Fox, B., Haldeman, B., Ostrander, C.D., Kaifu, T., Chabannon, C., Moretta, A., West, R., Xu, W., Vivier, E. & Levin, S.D. The B7 family member B7-H6 is a tumor cell ligand for the activating natural killer cell receptor NKp30 in humans. *J Exp Med* **206**, 1495-503 (2009).
48. Li, Y., Wang, Q. & Mariuzza, R.A. Structure of the human activating natural cytotoxicity receptor NKp30 bound to its tumor cell ligand B7-H6. *J Exp Med* **208**, 703-14 (2011).
49. Cosman, D., Mullberg, J., Sutherland, C.L., Chin, W., Armitage, R., Fanslow, W., Kubin, M. & Chalupny, N.J. ULBPs, novel MHC class I-related molecules, bind to CMV glycoprotein UL16 and stimulate NK cytotoxicity through the NKG2D receptor. *Immunity* **14**, 123-33 (2001).
50. Perussia, B., Trinchieri, G., Jackson, A., Warner, N.L., Faust, J., Rumpold, H., Kraft, D. & Lanier, L.L. The Fc receptor for IgG on human natural killer cells: phenotypic, functional, and comparative studies with monoclonal antibodies. *J Immunol* **133**, 180-9 (1984).
51. Veeramani, S., Wang, S.Y., Dahle, C., Blackwell, S., Jacobus, L., Knutson, T., Button, A., Link, B.K. & Weiner, G.J. Rituximab infusion induces NK activation in lymphoma patients with the high-affinity CD16 polymorphism. *Blood* **118**, 3347-9 (2011).
52. Cartron, G., Dacheux, L., Salles, G., Solal-Celigny, P., Bardos, P., Colombat, P. & Watier, H. Therapeutic activity of humanized anti-CD20 monoclonal antibody and polymorphism in IgG Fc receptor FcgammaRIIIa gene. *Blood* **99**, 754-8 (2002).
53. Kubota, A., Kubota, S., Lohwasser, S., Mager, D.L. & Takei, F. Diversity of NK cell receptor repertoire in adult and neonatal mice. *J Immunol* **163**, 212-6 (1999).
54. Valiante, N.M., Uhrberg, M., Shilling, H.G., Lienert-Weidenbach, K., Arnett, K.L., D'Andrea, A., Phillips, J.H., Lanier, L.L. & Parham, P. Functionally and structurally distinct NK cell receptor repertoires in the peripheral blood of two human donors. *Immunity* **7**, 739-51 (1997).
55. Raulet, D.H., Held, W., Correa, I., Dorfman, J.R., Wu, M.F. & Corral, L. Specificity, tolerance and developmental regulation of natural killer cells defined by expression of class I-specific Ly49 receptors. *Immunol Rev* **155**, 41-52 (1997).
56. Yokoyama, W.M. & Plougastel, B.F. Immune functions encoded by the natural killer gene complex. *Nat Rev Immunol* **3**, 304-16 (2003).
57. Raulet, D.H. & Vance, R.E. Self-tolerance of natural killer cells. *Nat Rev Immunol* **6**, 520-31 (2006).
58. Raulet, D.H. Missing self recognition and self tolerance of natural killer (NK) cells. *Semin Immunol* **18**, 145-50 (2006).
59. Kim, S., Poursine-Laurent, J., Truscott, S.M., Lybarger, L., Song, Y.J., Yang, L., French, A.R., Sunwoo, J.B., Lemieux, S., Hansen, T.H. & Yokoyama, W.M. Licensing of natural killer cells by host major histocompatibility complex class I molecules. *Nature* **436**, 709-13 (2005).
60. Anfossi, N., Andre, P., Guia, S., Falk, C.S., Roetynck, S., Stewart, C.A., Breso, V., Frassati, C., Reviron, D., Middleton, D., Romagne, F., Ugolini, S. & Vivier, E. Human NK cell education by inhibitory receptors for MHC class I. *Immunity* **25**, 331-42 (2006).
61. Hoglund, P., Glas, R., Menard, C., Kase, A., Johansson, M.H., Franksson, L., Lemmonier, F. & Karre, K. Beta2-microglobulin-deficient NK cells show increased sensitivity to MHC class I-mediated inhibition, but self tolerance does not depend upon target cell expression of H-2Kb and Db heavy chains. *Eur J Immunol* **28**, 370-8 (1998).
62. Liao, N.S., Bix, M., Zijlstra, M., Jaenisch, R. & Raulet, D. MHC class I deficiency: susceptibility to natural killer (NK) cells and impaired NK activity. *Science* **253**, 199-202 (1991).
63. Zimmer, J., Donato, L., Hanau, D., Cazenave, J.P., Tongio, M.M., Moretta, A. & de la Salle, H. Activity and phenotype of natural killer cells in peptide transporter (TAP)-deficient patients (type I bare lymphocyte syndrome). *J Exp Med* **187**, 117-22 (1998).
64. Vivier, E., Ugolini, S., Blaise, D., Chabannon, C. & Brossay, L. Targeting natural killer cells and natural killer T cells in cancer. *Nat Rev Immunol* **12**, 239-52 (2012).
65. Guia, S., Jaeger, B.N., Piatek, S., Mailfert, S., Trombik, T., Fenis, A., Chevrier, N., Walzer, T., Kerdiles, Y.M., Marguet, D., Vivier, E. & Ugolini, S. Confinement of activating receptors at the plasma membrane controls natural killer cell tolerance. *Sci Signal* **4**, ra21 (2011).
66. Joncker, N.T., Shifrin, N., Delebecque, F. & Raulet, D.H. Mature natural killer cells reset their responsiveness when exposed to an altered MHC environment. *J Exp Med* **207**, 2065-72 (2010).
67. Elliott, J.M., Wahle, J.A. & Yokoyama, W.M. MHC class I-deficient natural killer cells acquire a

licensed phenotype after transfer into an MHC class I-sufficient environment. *J Exp Med* **207**, 2073-9 (2010).
68. Long, E.O. Regulation of immune responses through inhibitory receptors. *Annu Rev Immunol* **17**, 875-904 (1999).
69. Valiante, N.M., Phillips, J.H., Lanier, L.L. & Parham, P. Killer cell inhibitory receptor recognition of human leukocyte antigen (HLA) class I blocks formation of a pp36/PLC-gamma signaling complex in human natural killer (NK) cells. *J Exp Med* **184**, 2243-50 (1996).
70. Binstadt, B.A., Brumbaugh, K.M., Dick, C.J., Scharenberg, A.M., Williams, B.L., Colonna, M., Lanier, L.L., Kinet, J.P., Abraham, R.T. & Leibson, P.J. Sequential involvement of Lck and SHP-1 with MHC-recognizing receptors on NK cells inhibits FcR-initiated tyrosine kinase activation. *Immunity* **5**, 629-38 (1996).
71. Binstadt, B.A., Billadeau, D.D., Jevremovic, D., Williams, B.L., Fang, N., Yi, T., Koretzky, G.A., Abraham, R.T. & Leibson, P.J. SLP-76 is a direct substrate of SHP-1 recruited to killer cell inhibitory receptors. *J Biol Chem* **273**, 27518-23 (1998).
72. Palmieri, G., Tullio, V., Zingoni, A., Piccoli, M., Frati, L., Lopez-Botet, M. & Santoni, A. CD94/NKG2-A inhibitory complex blocks CD16-triggered Syk and extracellular regulated kinase activation, leading to cytotoxic function of human NK cells. *J Immunol* **162**, 7181-8 (1999).
73. Stebbins, C.C., Watzl, C., Billadeau, D.D., Leibson, P.J., Burshtyn, D.N. & Long, E.O. Vav1 dephosphorylation by the tyrosine phosphatase SHP-1 as a mechanism for inhibition of cellular cytotoxicity. *Mol Cell Biol* **23**, 6291-9 (2003).
74. Anderson, P., Caligiuri, M., Ritz, J. & Schlossman, S.F. CD3-negative natural killer cells express zeta TCR as part of a novel molecular complex. *Nature* **341**, 159-62 (1989).
75. Hibbs, M.L., Selvaraj, P., Carpen, O., Springer, T.A., Kuster, H., Jouvin, M.H. & Kinet, J.P. Mechanisms for regulating expression of membrane isoforms of Fc gamma RIII (CD16). *Science* **246**, 1608-11 (1989).
76. Lanier, L.L., Corliss, B.C., Wu, J., Leong, C. & Phillips, J.H. Immunoreceptor DAP12 bearing a tyrosine-based activation motif is involved in activating NK cells. *Nature* **391**, 703-7 (1998).
77. Reth, M. Antigen receptor tail clue. *Nature* **338**, 383-4 (1989).
78. Colucci, F., Schweighoffer, E., Tomasello, E., Turner, M., Ortaldo, J.R., Vivier, E., Tybulewicz, V.L. & Di Santo, J.P. Natural cytotoxicity uncoupled from the Syk and ZAP-70 intracellular kinases. *Nat Immunol* **3**, 288-94 (2002).
79. Diefenbach, A., Tomasello, E., Lucas, M., Jamieson, A.M., Hsia, J.K., Vivier, E. & Raulet, D.H. Selective associations with signaling proteins determine stimulatory versus costimulatory activity of NKG2D. *Nat Immunol* **3**, 1142-9 (2002).
80. Gilfillan, S., Ho, E.L., Cella, M., Yokoyama, W.M. & Colonna, M. NKG2D recruits two distinct adapters to trigger NK cell activation and costimulation. *Nat Immunol* **3**, 1150-5 (2002).
81. Natarajan, K., Dimasi, N., Wang, J., Mariuzza, R.A. & Margulies, D.H. Structure and function of natural killer cell receptors: multiple molecular solutions to self, nonself discrimination. *Annu Rev Immunol* **20**, 853-85 (2002).
82. Wilson, M.J., Torkar, M., Haude, A., Milne, S., Jones, T., Sheer, D., Beck, S. & Trowsdale, J. Plasticity in the organization and sequences of human KIR/ILT gene families. *Proc Natl Acad Sci U S A* **97**, 4778-83 (2000).
83. Renedo, M., Arce, I., Montgomery, K., Roda-Navarro, P., Lee, E., Kucherlapati, R. & Fernandez-Ruiz, E. A sequence-ready physical map of the region containing the human natural killer gene complex on chromosome 12p12.3-p13.2. *Genomics* **65**, 129-36 (2000).
84. Yokoyama, W.M., Kehn, P.J., Cohen, D.I. & Shevach, E.M. Chromosomal location of the Ly-49 (A1, YE1/48) multigene family. Genetic association with the NK 1.1 antigen. *J Immunol* **145**, 2353-8 (1990).
85. Uhrberg, M., Valiante, N.M., Shum, B.P., Shilling, H.G., Lienert-Weidenbach, K., Corliss, B., Tyan, D., Lanier, L.L. & Parham, P. Human diversity in killer cell inhibitory receptor genes. *Immunity* **7**, 753-63 (1997).
86. Abi-Rached, L. & Parham, P. Natural selection drives recurrent formation of activating killer cell immunoglobulin-like receptor and Ly49 from inhibitory homologues. *J Exp Med* **201**, 1319-32 (2005).
87. Abi-Rached, L., Moesta, A.K., Rajalingam, R., Guethlein, L.A. & Parham, P. Human-specific evolution and adaptation led to major qualitative differences in the variable receptors of human and chimpanzee natural killer cells. *PLoS Genet* **6**, e1001192 (2010).
88. Barten, R., Torkar, M., Haude, A., Trowsdale, J. & Wilson, M.J. Divergent and convergent evolution of NK-cell receptors. *Trends Immunol* **22**, 52-7 (2001).
89. Kelley, J., Walter, L. & Trowsdale, J. Comparative genomics of natural killer cell receptor gene clusters. *PLoS Genet* **1**, 129-39 (2005).
90. Sambrook, J.G., Bashirova, A., Palmer, S., Sims, S., Trowsdale, J., Abi-Rached, L., Parham, P., Carrington, M. & Beck, S. Single haplotype analysis demonstrates rapid evolution of the

killer immunoglobulin-like receptor (KIR) loci in primates. *Genome Res* **15**, 25-35 (2005).
91. Mehta, I.K., Smith, H.R., Wang, J., Margulies, D.H. & Yokoyama, W.M. A "chimeric" C57l-derived Ly49 inhibitory receptor resembling the Ly49D activation receptor. *Cell Immunol* **209**, 29-41 (2001).
92. Trowsdale, J. Genetic and functional relationships between MHC and NK receptor genes. *Immunity* **15**, 363-74 (2001).
93. Hao, L. & Nei, M. Genomic organization and evolutionary analysis of Ly49 genes encoding the rodent natural killer cell receptors: rapid evolution by repeated gene duplication. *Immunogenetics* **56**, 343-54 (2004).
94. Yokoyama, W.M. Hybrid resistance and the Ly-49 family of natural killer cell receptors. *J Exp Med* **182**, 273-7 (1995).
95. Marsh, S.G., Parham, P., Dupont, B., Geraghty, D.E., Trowsdale, J., Middleton, D., Vilches, C., Carrington, M., Witt, C., Guethlein, L.A., Shilling, H., Garcia, C.A., Hsu, K.C. & Wain, H. Killer-cell immunoglobulin-like receptor (KIR) nomenclature report, 2002. *Immunogenetics* **55**, 220-6 (2003).
96. Goodridge, J.P., Witt, C.S., Christiansen, F.T. & Warren, H.S. KIR2DL4 (CD158d) genotype influences expression and function in NK cells. *J Immunol* **171**, 1768-74 (2003).
97. Stewart, C.A., Laugier-Anfossi, F., Vely, F., Saulquin, X., Riedmuller, J., Tisserant, A., Gauthier, L., Romagne, F., Ferracci, G., Arosa, F.A., Moretta, A., Sun, P.D., Ugolini, S. & Vivier, E. Recognition of peptide-MHC class I complexes by activating killer immunoglobulin-like receptors. *Proc Natl Acad Sci U S A* **102**, 13224-9 (2005).
98. Chewning, J.H., Gudme, C.N., Hsu, K.C., Selvakumar, A. & Dupont, B. KIR2DS1-positive NK cells mediate alloresponse against the C2 HLA-KIR ligand group in vitro. *J Immunol* **179**, 854-68 (2007).
99. Boyington, J.C., Motyka, S.A., Schuck, P., Brooks, A.G. & Sun, P.D. Crystal structure of an NK cell immunoglobulin-like receptor in complex with its class I MHC ligand. *Nature* **405**, 537-43 (2000).
100. Fan, Q.R., Long, E.O. & Wiley, D.C. Crystal structure of the human natural killer cell inhibitory receptor KIR2DL1-HLA-Cw4 complex. *Nat Immunol* **2**, 452-60 (2001).
101. Maenaka, K., Juji, T., Stuart, D.I. & Jones, E.Y. Crystal structure of the human p58 killer cell inhibitory receptor (KIR2DL3) specific for HLA-Cw3-related MHC class I. *Structure* **7**, 391-8 (1999).
102. Moesta, A.K., Norman, P.J., Yawata, M., Yawata, N., Gleimer, M. & Parham, P. Synergistic polymorphism at two positions distal to the ligand-binding site makes KIR2DL2 a stronger receptor for HLA-C than KIR2DL3. *J Immunol* **180**, 3969-79 (2008).
103. Boyington, J.C., Brooks, A.G. & Sun, P.D. Structure of killer cell immunoglobulin-like receptors and their recognition of the class I MHC molecules. *Immunol Rev* **181**, 66-78 (2001).
104. Reinherz, E.L., Tan, K., Tang, L., Kern, P., Liu, J., Xiong, Y., Hussey, R.E., Smolyar, A., Hare, B., Zhang, R., Joachimiak, A., Chang, H.C., Wagner, G. & Wang, J. The crystal structure of a T cell receptor in complex with peptide and MHC class II. *Science* **286**, 1913-21 (1999).
105. Hennecke, J., Carfi, A. & Wiley, D.C. Structure of a covalently stabilized complex of a human alphabeta T-cell receptor, influenza HA peptide and MHC class II molecule, HLA-DR1. *EMBO J* **19**, 5611-24 (2000).
106. Hansasuta, P., Dong, T., Thananchai, H., Weekes, M., Willberg, C., Aldemir, H., Rowland-Jones, S. & Braud, V.M. Recognition of HLA-A3 and HLA-A11 by KIR3DL2 is peptide-specific. *Eur J Immunol* **34**, 1673-9 (2004).
107. Malnati, M.S., Peruzzi, M., Parker, K.C., Biddison, W.E., Ciccone, E., Moretta, A. & Long, E.O. Peptide specificity in the recognition of MHC class I by natural killer cell clones. *Science* **267**, 1016-8 (1995).
108. Rajagopalan, S. & Long, E.O. The direct binding of a p58 killer cell inhibitory receptor to human histocompatibility leukocyte antigen (HLA)-Cw4 exhibits peptide selectivity. *J Exp Med* **185**, 1523-8 (1997).
109. Zappacosta, F., Borrego, F., Brooks, A.G., Parker, K.C. & Coligan, J.E. Peptides isolated from HLA-Cw*0304 confer different degrees of protection from natural killer cell-mediated lysis. *Proc Natl Acad Sci U S A* **94**, 6313-8 (1997).
110. Sharma, D., Bastard, K., Guethlein, L.A., Norman, P.J., Yawata, N., Yawata, M., Pando, M., Thananchai, H., Dong, T., Rowland-Jones, S., Brodsky, F.M. & Parham, P. Dimorphic motifs in D0 and D1+D2 domains of killer cell Ig-like receptor 3DL1 combine to form receptors with high, moderate, and no avidity for the complex of a peptide derived from HIV and HLA-A*2402. *J Immunol* **183**, 4569-82 (2009).
111. Wilhelm, B.T., Gagnier, L. & Mager, D.L. Sequence analysis of the ly49 cluster in C57BL/6 mice: a rapidly evolving multigene family in the immune system. *Genomics* **80**, 646-61 (2002).
112. Makrigiannis, A.P., Pau, A.T., Saleh, A., Winkler-Pickett, R., Ortaldo, J.R. & Anderson, S.K. Class I

MHC-binding characteristics of the 129/J Ly49 repertoire. *J Immunol* **166**, 5034-43 (2001).
113.Anderson, S.K., Ortaldo, J.R. & McVicar, D.W. The ever-expanding Ly49 gene family: repertoire and signaling. *Immunol Rev* **181**, 79-89 (2001).
114.Zelensky, A.N. & Gready, J.E. The C-type lectin-like domain superfamily. *FEBS J* **272**, 6179-217 (2005).
115.Tormo, J., Natarajan, K., Margulies, D.H. & Mariuzza, R.A. Crystal structure of a lectin-like natural killer cell receptor bound to its MHC class I ligand. *Nature* **402**, 623-31 (1999).
116.Dam, J., Guan, R., Natarajan, K., Dimasi, N., Chlewicki, L.K., Kranz, D.M., Schuck, P., Margulies, D.H. & Mariuzza, R.A. Variable MHC class I engagement by Ly49 natural killer cell receptors demonstrated by the crystal structure of Ly49C bound to H-2K(b). *Nat Immunol* **4**, 1213-22 (2003).
117.Dam, J., Baber, J., Grishaev, A., Malchiodi, E.L., Schuck, P., Bax, A. & Mariuzza, R.A. Variable dimerization of the Ly49A natural killer cell receptor results in differential engagement of its MHC class I ligand. *J Mol Biol* **362**, 102-13 (2006).
118.Held, W. & Mariuzza, R.A. Cis interactions of immunoreceptors with MHC and non-MHC ligands. *Nat Rev Immunol* **8**, 269-78 (2008).
119.Doucey, M.A., Scarpellino, L., Zimmer, J., Guillaume, P., Luescher, I.F., Bron, C. & Held, W. Cis association of Ly49A with MHC class I restricts natural killer cell inhibition. *Nat Immunol* **5**, 328-36 (2004).
120.Masuda, A., Nakamura, A., Maeda, T., Sakamoto, Y. & Takai, T. Cis binding between inhibitory receptors and MHC class I can regulate mast cell activation. *J Exp Med* **204**, 907-20 (2007).
121.Scarpellino, L., Oeschger, F., Guillaume, P., Coudert, J.D., Levy, F., Leclercq, G. & Held, W. Interactions of Ly49 family receptors with MHC class I ligands in trans and cis. *J Immunol* **178**, 1277-84 (2007).
122.Back, J., Chalifour, A., Scarpellino, L. & Held, W. Stable masking by H-2Dd cis ligand limits Ly49A relocalization to the site of NK cell/target cell contact. *Proc Natl Acad Sci U S A* **104**, 3978-83 (2007).
123.Chalifour, A., Scarpellino, L., Back, J., Brodin, P., Devevre, E., Gros, F., Levy, F., Leclercq, G., Hoglund, P., Beermann, F. & Held, W. A Role for cis Interaction between the Inhibitory Ly49A receptor and MHC class I for natural killer cell education. *Immunity* **30**, 337-47 (2009).
124.Back, J., Malchiodi, E.L., Cho, S., Scarpellino, L., Schneider, P., Kerzic, M.C., Mariuzza, R.A. & Held, W. Distinct conformations of Ly49 natural killer cell receptors mediate MHC class I recognition in trans and cis. *Immunity* **31**, 598-608 (2009).
125.Correa, I. & Raulet, D.H. Binding of diverse peptides to MHC class I molecules inhibits target cell lysis by activated natural killer cells. *Immunity* **2**, 61-71 (1995).
126. Orihuela, M., Margulies, D.H. & Yokoyama, W.M. The natural killer cell receptor Ly-49A recognizes a peptide-induced conformational determinant on its major histocompatibility complex class I ligand. *Proc Natl Acad Sci U S A* **93**, 11792-7 (1996).
127.Silver, E.T., Gong, D.E., Chang, C.S., Amrani, A., Santamaria, P. & Kane, K.P. Ly-49P activates NK-mediated lysis by recognizing H-2Dd. *J Immunol* **165**, 1771-81 (2000).
128.Silver, E.T., Gong, D., Hazes, B. & Kane, K.P. Ly-49W, an activating receptor of nonobese diabetic mice with close homology to the inhibitory receptor Ly-49G, recognizes H-2D(k) and H-2D(d). *J Immunol* **166**, 2333-41 (2001).
129.Aramburu, J., Balboa, M.A., Ramirez, A., Silva, A., Acevedo, A., Sanchez-Madrid, F., De Landazuri, M.O. & Lopez-Botet, M. A novel functional cell surface dimer (Kp43) expressed by natural killer cells and T cell receptor-gamma/delta+ T lymphocytes. I. Inhibition of the IL-2-dependent proliferation by anti-Kp43 monoclonal antibody. *J Immunol* **144**, 3238-47 (1990).
130.Toyama-Sorimachi, N., Taguchi, Y., Yagita, H., Kitamura, F., Kawasaki, A., Koyasu, S. & Karasuyama, H. Mouse CD94 participates in Qa-1-mediated self recognition by NK cells and delivers inhibitory signals independent of Ly-49. *J Immunol* **166**, 3771-9 (2001).
131.Chang, C., Rodriguez, A., Carretero, M., Lopez-Botet, M., Phillips, J.H. & Lanier, L.L. Molecular characterization of human CD94: a type II membrane glycoprotein related to the C-type lectin superfamily. *Eur J Immunol* **25**, 2433-7 (1995).
132.Lanier, L.L., Corliss, B., Wu, J. & Phillips, J.H. Association of DAP12 with activating CD94/NKG2C NK cell receptors. *Immunity* **8**, 693-701 (1998).
133.Le Drean, E., Vely, F., Olcese, L., Cambiaggi, A., Guia, S., Krystal, G., Gervois, N., Moretta, A., Jotereau, F. & Vivier, E. Inhibition of antigen-induced T cell response and antibody-induced NK cell cytotoxicity by NKG2A: association of NKG2A with SHP-1 and SHP-2 protein-tyrosine phosphatases. *Eur J Immunol* **28**, 264-76 (1998).
134.Olcese, L., Cambiaggi, A., Semenzato, G., Bottino, C., Moretta, A. & Vivier, E. Human killer cell activatory receptors for MHC class I molecules are included in a multimeric complex

expressed by natural killer cells. *J Immunol* **158**, 5083-6 (1997).
135.Soderstrom, K., Corliss, B., Lanier, L.L. & Phillips, J.H. CD94/NKG2 is the predominant inhibitory receptor involved in recognition of HLA-G by decidual and peripheral blood NK cells. *J Immunol* **159**, 1072-5 (1997).
136.Brooks, A.G., Posch, P.E., Scorzelli, C.J., Borrego, F. & Coligan, J.E. NKG2A complexed with CD94 defines a novel inhibitory natural killer cell receptor. *J Exp Med* **185**, 795-800 (1997).
137.Posch, P.E., Borrego, F., Brooks, A.G. & Coligan, J.E. HLA-E is the ligand for the natural killer cell CD94/NKG2 receptors. *J Biomed Sci* **5**, 321-31 (1998).
138.Vance, R.E., Kraft, J.R., Altman, J.D., Jensen, P.E. & Raulet, D.H. Mouse CD94/NKG2A is a natural killer cell receptor for the nonclassical major histocompatibility complex (MHC) class I molecule Qa-1(b). *J Exp Med* **188**, 1841-8 (1998).
139.Braud, V., Jones, E.Y. & McMichael, A. The human major histocompatibility complex class Ib molecule HLA-E binds signal sequence-derived peptides with primary anchor residues at positions 2 and 9. *Eur J Immunol* **27**, 1164-9 (1997).
140.Borrego, F., Ulbrecht, M., Weiss, E.H., Coligan, J.E. & Brooks, A.G. Recognition of human histocompatibility leukocyte antigen (HLA)-E complexed with HLA class I signal sequence-derived peptides by CD94/NKG2 confers protection from natural killer cell-mediated lysis. *J Exp Med* **187**, 813-8 (1998).
141.Kurepa, Z., Hasemann, C.A. & Forman, J. Qa-1b binds conserved class I leader peptides derived from several mammalian species. *J Exp Med* **188**, 973-8 (1998).
142.Boyington, J.C., Riaz, A.N., Patamawenu, A., Coligan, J.E., Brooks, A.G. & Sun, P.D. Structure of CD94 reveals a novel C-type lectin fold: implications for the NK cell-associated CD94/NKG2 receptors. *Immunity* **10**, 75-82 (1999).
143.Sullivan, L.C., Clements, C.S., Beddoe, T., Johnson, D., Hoare, H.L., Lin, J., Huyton, T., Hopkins, E.J., Reid, H.H., Wilce, M.C., Kabat, J., Borrego, F., Coligan, J.E., Rossjohn, J. & Brooks, A.G. The heterodimeric assembly of the CD94-NKG2 receptor family and implications for human leukocyte antigen-E recognition. *Immunity* **27**, 900-11 (2007).
144.Petrie, E.J., Clements, C.S., Lin, J., Sullivan, L.C., Johnson, D., Huyton, T., Heroux, A., Hoare, H.L., Beddoe, T., Reid, H.H., Wilce, M.C., Brooks, A.G. & Rossjohn, J. CD94-NKG2A recognition of human leukocyte antigen (HLA)-E bound to an HLA class I leader sequence. *J Exp Med* **205**, 725-35 (2008).
145.Kaiser, B.K., Pizarro, J.C., Kerns, J. & Strong, R.K. Structural basis for NKG2A/CD94 recognition of HLA-E. *Proc Natl Acad Sci U S A* **105**, 6696-701 (2008).
146.Houchins, J.P., Yabe, T., McSherry, C. & Bach, F.H. DNA sequence analysis of NKG2, a family of related cDNA clones encoding type II integral membrane proteins on human natural killer cells. *J Exp Med* **173**, 1017-20 (1991).
147.Bahram, S., Bresnahan, M., Geraghty, D.E. & Spies, T. A second lineage of mammalian major histocompatibility complex class I genes. *Proc Natl Acad Sci U S A* **91**, 6259-63 (1994).
148. Bahram, S. & Spies, T. Nucleotide sequence of a human MHC class I MICB cDNA. *Immunogenetics* **43**, 230-3 (1996).
149.Groh, V., Rhinehart, R., Secrist, H., Bauer, S., Grabstein, K.H. & Spies, T. Broad tumor-associated expression and recognition by tumor-derived gamma delta T cells of MICA and MICB. *Proc Natl Acad Sci U S A* **96**, 6879-84 (1999).
150.Groh, V., Steinle, A., Bauer, S. & Spies, T. Recognition of stress-induced MHC molecules by intestinal epithelial gammadelta T cells. *Science* **279**, 1737-40 (1998).
151.Cerwenka, A., Bakker, A.B., McClanahan, T., Wagner, J., Wu, J., Phillips, J.H. & Lanier, L.L. Retinoic acid early inducible genes define a ligand family for the activating NKG2D receptor in mice. *Immunity* **12**, 721-7 (2000).
152.Diefenbach, A., Jamieson, A.M., Liu, S.D., Shastri, N. & Raulet, D.H. Ligands for the murine NKG2D receptor: expression by tumor cells and activation of NK cells and macrophages. *Nat Immunol* **1**, 119-26 (2000).
153.Groh, V., Wu, J., Yee, C. & Spies, T. Tumour-derived soluble MIC ligands impair expression of NKG2D and T-cell activation. *Nature* **419**, 734-8 (2002).
154.Doubrovina, E.S., Doubrovin, M.M., Vider, E., Sisson, R.B., O'Reilly, R.J., Dupont, B. & Vyas, Y.M. Evasion from NK cell immunity by MHC class I chain-related molecules expressing colon adenocarcinoma. *J Immunol* **171**, 6891-9 (2003).
155.Salih, H.R., Rammensee, H.G. & Steinle, A. Cutting edge: down-regulation of MICA on human tumors by proteolytic shedding. *J Immunol* **169**, 4098-102 (2002).
156.Li, P., Morris, D.L., Willcox, B.E., Steinle, A., Spies, T. & Strong, R.K. Complex structure of the activating immunoreceptor NKG2D and its MHC class I-like ligand MICA. *Nat Immunol* **2**, 443-51 (2001).

157.Li, P., McDermott, G. & Strong, R.K. Crystal structures of RAE-1beta and its complex with the activating immunoreceptor NKG2D. *Immunity* **16**, 77-86 (2002).
158.Plougastel, B., Matsumoto, K., Dubbelde, C. & Yokoyama, W.M. Analysis of a 1-Mb BAC contig overlapping the mouse Nkrp1 cluster of genes: cloning of three new Nkrp1 members, Nkrp1d, Nkrp1e, and Nkrp1f. *Immunogenetics* **53**, 592-8 (2001).
159.Chen, P., Belanger, S., Aguilar, O.A., Zhang, Q., St-Laurent, A., Rahim, M.M., Makrigiannis, A.P. & Carlyle, J.R. Analysis of the mouse 129-strain Nkrp1-Clr gene cluster reveals conservation of genomic organization and functional receptor-ligand interactions despite significant allelic polymorphism. *Immunogenetics* **63**, 627-40 (2011).
160.Zhang, Q., Rahim, M.M., Allan, D.S., Tu, M.M., Belanger, S., Abou-Samra, E., Ma, J., Sekhon, H.S., Fairhead, T., Zein, H.S., Carlyle, J.R., Anderson, S.K. & Makrigiannis, A.P. Mouse Nkrp1-Clr gene cluster sequence and expression analyses reveal conservation of tissue-specific MHC-independent immunosurveillance. *PLoS One* **7**, e50561 (2012).
161.Kveberg, L., Dai, K.Z., Westgaard, I.H., Daws, M.R., Fossum, S., Naper, C. & Vaage, J.T. Two major groups of rat NKR-P1 receptors can be distinguished based on chromosomal localization, phylogenetic analysis and Clr ligand binding. *Eur J Immunol* **39**, 541-51 (2009).
162.Lanier, L.L., Chang, C. & Phillips, J.H. Human NKR-P1A. A disulfide-linked homodimer of the C-type lectin superfamily expressed by a subset of NK and T lymphocytes. *J Immunol* **153**, 2417-28 (1994).
163.Kveberg, L., Dai, K.Z., Inngjerdingen, M., Brooks, C.G., Fossum, S. & Vaage, J.T. Phylogenetic and functional conservation of the NKR-P1F and NKR-P1G receptors in rat and mouse. *Immunogenetics* **63**, 429-36 (2011).
164.Aust, J.G., Gays, F., Mickiewicz, K.M., Buchanan, E. & Brooks, C.G. The expression and function of the NKRP1 receptor family in C57BL/6 mice. *J Immunol* **183**, 106-16 (2009).
165.Ryan, J.C., Turck, J., Niemi, E.C., Yokoyama, W.M. & Seaman, W.E. Molecular cloning of the NK1.1 antigen, a member of the NKR-P1 family of natural killer cell activation molecules. *J Immunol* **149**, 1631-5 (1992).
166.Arase, H., Arase, N. & Saito, T. Interferon gamma production by natural killer (NK) cells and NK1.1+ T cells upon NKR-P1 cross-linking. *J Exp Med* **183**, 2391-6 (1996).
167.Arase, N., Arase, H., Park, S.Y., Ohno, H., Ra, C. & Saito, T. Association with FcRgamma is essential for activation signal through NKR-P1 (CD161) in natural killer (NK) cells and NK1.1+ T cells. *J Exp Med* **186**, 1957-63 (1997).
168.Glimcher, L., Shen, F.W. & Cantor, H. Identification of a cell-surface antigen selectively expressed on the natural killer cell. *J Exp Med* **145**, 1-9 (1977).
169.Koo, G.C. & Peppard, J.R. Establishment of monoclonal anti-Nk-1.1 antibody. *Hybridoma* **3**, 301-3 (1984).
170.Carlyle, J.R., Martin, A., Mehra, A., Attisano, L., Tsui, F.W. & Zuniga-Pflucker, J.C. Mouse NKR-P1B, a novel NK1.1 antigen with inhibitory function. *J Immunol* **162**, 5917-23 (1999).
171.Kung, S.K., Su, R.C., Shannon, J. & Miller, R.G. The NKR-P1B gene product is an inhibitory receptor on SJL/J NK cells. *J Immunol* **162**, 5876-87 (1999).
172.Liu, J., Morris, M.A., Nguyen, P., George, T.C., Koulich, E., Lai, W.C., Schatzle, J.D., Kumar, V. & Bennett, M. Ly49I NK cell receptor transgene inhibition of rejection of H2b mouse bone marrow transplants. *J Immunol* **164**, 1793-9 (2000).
173.Carlyle, J.R., Mesci, A., Ljutic, B., Belanger, S., Tai, L.H., Rousselle, E., Troke, A.D., Proteau, M.F. & Makrigiannis, A.P. Molecular and genetic basis for strain-dependent NK1.1 alloreactivity of mouse NK cells. *J Immunol* **176**, 7511-24 (2006).
174.Bezouska, K., Vlahas, G., Horvath, O., Jinochova, G., Fiserova, A., Giorda, R., Chambers, W.H., Feizi, T. & Pospisil, M. Rat natural killer cell antigen, NKR-P1, related to C-type animal lectins is a carbohydrate-binding protein. *J Biol Chem* **269**, 16945-52 (1994).
175.Bezouska, K., Yuen, C.T., O'Brien, J., Childs, R.A., Chai, W., Lawson, A.M., Drbal, K., Fiserova, A., Pospisil, M. & Feizi, T. Oligosaccharide ligands for NKR-P1 protein activate NK cells and cytotoxicity. *Nature* **372**, 150-7 (1994).
176.Kogelberg, H., Montero, E., Bay, S., Lawson, A.M. & Feizi, T. Re-evaluation of monosaccharide binding property of recombinant soluble carbohydrate recognition domain of the natural killer cell receptor NKR-P1A. *J Biol Chem* **274**, 30335-6 (1999).
177.Kogelberg, H., Frenkiel, T.A., Birdsall, B., Chai, W. & Muskett, F.W. Binding of sucrose octasulphate to the C-type lectin-like domain of the recombinant natural killer cell receptor NKR-P1A observed by NMR spectroscopy. *Chembiochem* **3**, 1072-7 (2002).
178.Bezouska, K., Sklenar, J., Dvorakova, J., Havlicek, V., Pospisil, M., Thiem, J. & Kren, V. NKR-P1A protein, an activating receptor of rat natural killer cells, binds to the chitobiose core of uncompletely glycosylated N-linked glycans,

and to linear chitooligomers. *Biochem Biophys Res Commun* **238**, 149-53 (1997).
179.Krist, P., Herkommerova-Rajnochova, E., Rauvolfova, J., Semenuk, T., Vavruskova, P., Pavlicek, J., Bezouska, K., Petrus, L. & Kren, V. Toward an optimal oligosaccharide ligand for rat natural killer cell activation receptor NKR-P1. *Biochem Bioph Res Co* **287**, 11-20 (2001).
180.Sedmera, P., Prikrylova, V., Bezouska, K., Rajnochova, E., Thiem, J. & Kren, V. Preparation of ManNAc containing chitooligomers by isomerisation and their binding to NKR-P1 protein. *J Carbohyd Chem* **17**, 1351-1357 (1998).
181.Kren, V., Dvorakova, J., Gambert, U., Sedmera, P., Havlicek, V., Thiem, J. & Bezouska, K. beta-Glucosylation of chitooligomers by galactosyltransferase. *Carbohyd Res* **305**, 517-523 (1997).
182.Vannucci, L., Fiserova, A., Sadalapure, K., Lindhorst, T.K., Kuldova, M., Rossmann, P., Horvath, O., Kren, V., Krist, P., Bezouska, K., Luptovcova, M., Mosca, F. & Pospisil, M. Effects of N-acetyl-glucosamine-coated glycodendrimers as biological modulators in the B16F10 melanoma model in vivo. *Int J Oncol* **23**, 285-296 (2003).
183.Attolino, E., Bonaccorsi, F., Catelani, G., D'Andrea, F., Krenek, K., Bezouska, K. & Kren, V. Improved preparation of beta-D-ManNAc-(1 -> 4)-D-Glc and beta-D-TalNAc-(1 -> 4)-D-Glc disaccharides and evaluation of their activating properties on the natural killer cells NKR-P1 and CD69 receptors. *J Carbohyd Chem* **27**, 156-171 (2008).
184.Catelani, G., D'Andrea, F., Griselli, A., Guazzelli, L., Nemcova, P., Bezouska, K., Krenek, K. & Kren, V. Deoxynojirimycin and its hexosaminyl derivatives bind to natural killer cell receptors rNKR-P1A and hCD69. *Bioorg Med Chem Lett* **20**, 4645-8 (2010).
185.Bojarova, P., Krenek, K., Kuzma, M., Petraskova, L., Bezouska, K., Namdjou, D.J., Elling, L. & Kren, V. N-acetylhexosamine triad in one molecule: Chemoenzymatic introduction of 2-acetamido-2-deoxy-beta-D-galactopyranosyluronic acid residue into a complex oligosaccharide. *J Mol Catal B-Enzym* **50**, 69-73 (2008).
186.Fialova, P., Namdjou, D.J., Ettrich, R., Prikrylova, W., Rauvolfova, J., Krenek, K., Kuzma, M., Elling, L., Bezouska, K. & Kren, V. Combined application of galactose oxidase and beta-N-acetylhexosaminidase in the synthesis of complex immunoactive N-acetyl-D-galactosaminides. *Adv Synth Catal* **347**, 997-1006 (2005).
187.Bojarova, P., Krenek, K., Wetjen, K., Adamiak, K., Pelantova, H., Bezouska, K., Elling, L. & Kren, V. Synthesis of LacdiNAc-terminated glycoconjugates by mutant galactosyltransferase--a way to new glycodrugs and materials. *Glycobiology* **19**, 509-17 (2009).
188.Drozdova, A., Bojarova, P., Krenek, K., Weignerova, L., Henssen, B., Elling, L., Christensen, H., Jensen, H.H., Pelantova, H., Kuzma, M., Bezouska, K., Krupova, M., Adamek, D., Slamova, K. & Kren, V. Enzymatic synthesis of dimeric glycomimetic ligands of NK cell activation receptors. *Carbohydr Res* **346**, 1599-609 (2011).
189.Carlyle, J.R., Jamieson, A.M., Gasser, S., Clingan, C.S., Arase, H. & Raulet, D.H. Missing self-recognition of Ocil/Clr-b by inhibitory NKR-P1 natural killer cell receptors. *Proc Natl Acad Sci U S A* **101**, 3527-32 (2004).
190.Plougastel, B., Dubbelde, C. & Yokoyama, W.M. Cloning of Clr, a new family of lectin-like genes localized between mouse Nkrp1a and Cd69. *Immunogenetics* **53**, 209-14 (2001).
191.Zhou, H., Kartsogiannis, V., Hu, Y.S., Elliott, J., Quinn, J.M., McKinstry, W.J., Gillespie, M.T. & Ng, K.W. A novel osteoblast-derived C-type lectin that inhibits osteoclast formation. *J Biol Chem* **276**, 14916-23 (2001).
192.Zhou, H., Kartsogiannis, V., Quinn, J.M., Ly, C., Gange, C., Elliott, J., Ng, K.W. & Gillespie, M.T. Osteoclast inhibitory lectin, a family of new osteoclast inhibitors. *J Biol Chem* **277**, 48808-15 (2002).
193.Rosen, D.B., Bettadapura, J., Alsharifi, M., Mathew, P.A., Warren, H.S. & Lanier, L.L. Cutting edge: lectin-like transcript-1 is a ligand for the inhibitory human NKR-P1A receptor. *J Immunol* **175**, 7796-9 (2005).
194.Aldemir, H., Prod'homme, V., Dumaurier, M.J., Retiere, C., Poupon, G., Cazareth, J., Bihl, F. & Braud, V.M. Cutting edge: lectin-like transcript 1 is a ligand for the CD161 receptor. *J Immunol* **175**, 7791-5 (2005).
195.Poggi, A., Costa, P., Zocchi, M.R. & Moretta, L. NKRP1A molecule is involved in transendothelial migration of CD4+ human T lymphocytes. *Immunol Lett* **57**, 121-3 (1997).
196.Exley, M., Porcelli, S., Furman, M., Garcia, J. & Balk, S. CD161 (NKR-P1A) costimulation of CD1d-dependent activation of human T cells expressing invariant V alpha 24 J alpha Q T cell receptor alpha chains. *J Exp Med* **188**, 867-76 (1998).
197.Mathew, P.A., Chuang, S.S., Vaidya, S.V., Kumaresan, P.R., Boles, K.S. & Pham, H.T. The LLT1 receptor induces IFN-gamma production by human natural killer cells. *Mol Immunol* **40**, 1157-63 (2004).
198.Chambers, W.H., Vujanovic, N.L., DeLeo, A.B., Olszowy, M.W., Herberman, R.B. & Hiserodt, J.C.

Monoclonal antibody to a triggering structure expressed on rat natural killer cells and adherent lymphokine-activated killer cells. *J Exp Med* **169**, 1373-89 (1989).

199.Giorda, R., Rudert, W.A., Vavassori, C., Chambers, W.H., Hiserodt, J.C. & Trucco, M. NKR-P1, a signal transduction molecule on natural killer cells. *Science* **249**, 1298-300 (1990).

200.Rozbesky, D., Kavan, D., Chmelik, J., Novak, P., Vanek, O. & Bezouska, K. High-level expression of soluble form of mouse natural killer cell receptor NKR-P1C(B6) in Escherichia coli. *Protein Expr Purif* **77**, 178-84 (2011).

201.Ljutic, B., Carlyle, J.R. & Zuniga-Pflucker, J.C. Identification of upstream cis-acting regulatory elements controlling lineage-specific expression of the mouse NK cell activation receptor, NKR-P1C. *J Biol Chem* **278**, 31909-17 (2003).

202.Kogelberg, H., Lawson, A.M., Muskett, F.W., Carruthers, R.A. & Feizi, T. Expression in Escherichia coli, folding in vitro, and characterization of the carbohydrate recognition domain of the natural killer cell receptor NKR-P1A. *Protein Expr Purif* **20**, 10-20 (2000).

203.Vanek, O., Nalezkova, M., Kavan, D., Borovickova, I., Pompach, P., Novak, P., Kumar, V., Vannucci, L., Hudecek, J., Hofbauerova, K., Kopecky, V., Jr., Brynda, J., Kolenko, P., Dohnalek, J., Kaderavek, P., Chmelik, J., Gorcik, L., Zidek, L., Sklenar, V. & Bezouska, K. Soluble recombinant CD69 receptors optimized to have an exceptional physical and chemical stability display prolonged circulation and remain intact in the blood of mice. *FEBS J* **275**, 5589-606 (2008).

204.Pompach, P., Man, P., Kavan, D., Hofbauerova, K., Kumar, V., Bezouska, K., Havlicek, V. & Novak, P. Modified electrophoretic and digestion conditions allow a simplified mass spectrometric evaluation of disulfide bonds. *J Mass Spectrom* **44**, 1571-8 (2009).

205.Sovova, Z., Kopecky, V., Jr., Pazderka, T., Hofbauerova, K., Rozbesky, D., Vanek, O., Bezouska, K. & Ettrich, R. Structural analysis of natural killer cell receptor protein 1 (NKR-P1) extracellular domains suggests a conserved long loop region involved in ligand specificity. *J Mol Model* **17**, 1353-70 (2011).

206.Kolenko, P., Rozbesky, D., Vanek, O., Kopecky, V., Jr., Hofbauerova, K., Novak, P., Pompach, P., Hasek, J., Skalova, T., Bezouska, K. & Dohnalek, J. Molecular architecture of mouse activating NKR-P1 receptors. *J Struct Biol* **175**, 434-41 (2011).

207.Feinberg, H., Park-Snyder, S., Kolatkar, A.R., Heise, C.T., Taylor, M.E. & Weis, W.I. Structure of a C-type carbohydrate recognition domain from the macrophage mannose receptor. *J Biol Chem* **275**, 21539-48 (2000).

208.Maita, N., Nishio, K., Nishimoto, E., Matsui, T., Shikamoto, Y., Morita, T., Sadler, J.E. & Mizuno, H. Crystal structure of von Willebrand factor A1 domain complexed with snake venom, bitiscetin: insight into glycoprotein Ibalpha binding mechanism induced by snake venom proteins. *J Biol Chem* **278**, 37777-81 (2003).

209.Mizuno, H., Fujimoto, Z., Koizumi, M., Kano, H., Atoda, H. & Morita, T. Structure of coagulation factors IX/X-binding protein, a heterodimer of C-type lectin domains. *Nat Struct Biol* **4**, 438-41 (1997).

210.Mizuno, H., Fujimoto, Z., Atoda, H. & Morita, T. Crystal structure of an anticoagulant protein in complex with the Gla domain of factor X. *Proc Natl Acad Sci U S A* **98**, 7230-4 (2001).

211.Morita, T. Structures and functions of snake venom CLPs (C-type lectin-like proteins) with anticoagulant-, procoagulant-, and platelet-modulating activities. *Toxicon* **45**, 1099-114 (2005).

212.Watson, A.A., Lebedev, A.A., Hall, B.A., Fenton-May, A.E., Vagin, A.A., Dejnirattisai, W., Felce, J., Mongkolsapaya, J., Palma, A.S., Liu, Y., Feizi, T., Screaton, G.R., Murshudov, G.N. & O'Callaghan, C.A. Structural flexibility of the macrophage dengue virus receptor CLEC5A: implications for ligand binding and signaling. *J Biol Chem* **286**, 24208-18 (2011).

213.Kolenko, P., Rozbesky, D., Vanek, O., Bezouska, K., Hasek, J. & Dohnalek, J. Structure of the H107R variant of the extracellular domain of mouse NKR-P1A at 2.3 A resolution. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **67**, 1519-23 (2011).

214.Rozbesky, D., Man, P., Kavan, D., Chmelik, J., Cerny, J., Bezouska, K. & Novak, P. Chemical Cross-Linking and H/D Exchange for Fast Refinement of Protein Crystal Structure. *Anal Chem* **84**, 867-870 (2012).

215.Rozbesky, D., Sovova, Z., Marcoux, J., Man, P., Ettrich, R., Robinson, C.V. & Novak, P. Structural model of lymphocyte receptor NKR-P1C revealed by mass spectrometry and molecular modeling. *Anal Chem* **85**, 1597-604 (2013).

PAPER I

Rozbesky, D., Kavan, D., Chmelik, J., Novak, P., Vanek, O., Bezouska, K

High-level expression of soluble form of mouse natural killer cell receptor NKR-P1C(B6) in Escherichia coli.

Protein Expr. Purif. 77, 178-84 (2011)

My contribution to the publication: *research performing (cloning, protein expression, refolding in vitro, purification, expression and refolding optimi-zation, protein characterization), data collection, data analysis and interpretation, manuscript writing*

Protein Expression and Purification 77 (2011) 178–184

Contents lists available at ScienceDirect

Protein Expression and Purification

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

High-level expression of soluble form of mouse natural killer cell receptor NKR-P1C(B6) in *Escherichia coli*

Daniel Rozbeský [a,b], Daniel Kavan [b], Josef Chmelík [b], Petr Novák [a,b], Ondřej Vaněk [a,b], Karel Bezouška [a,b,*]

[a] Department of Biochemistry, Faculty of Science, Charles University in Prague, Hlavova 8, CZ-12840 Prague 2, Czech Republic
[b] Institute of Microbiology, v.v.i., Academy of Sciences of the Czech Republic, Vídeňská 1083, CZ-14220 Prague 4, Czech Republic

ARTICLE INFO

Article history:
Received 26 December 2010
and in revised form 25 January 2011
Available online 1 February 2011

Keywords:
Natural killer cell
NKR-P1C receptor
NK1.1 antigen
C-type lectin domain
Refolding

ABSTRACT

Mouse NKR-P1C(B6) receptor corresponding to NK1.1 alloantigen is one of the most widespread surface markers of mouse NK and NKT cells in C57BL/6 mice detected by monoclonal antibody PK136. Although functional studies revealed the ability of this receptor to activate both natural killing and production of cytokines upon antibody crosslinking, the ligand for NKR-P1C(B6) remains unknown. In order to initiate ligand identification, structural studies, and epitope mapping experiments, we developed a simple and efficient expression and purification protocol allowing to produce large amounts of pure soluble monomeric mouse NKR-P1C(B6). Our protein encompassed approximately half of the stalk region and the entire C-terminal globular ligand binding domain. The identity of protein that was devoid of N-terminal initiation methionine and had all three expected disulfides closed was confirmed using high resolution ion cyclotron resonance mass spectrometry. Protein produced into inclusion bodies in *Escherichia coli* was efficiently refolded into a unique three dimensional structure as confirmed by NMR using $^{1}H–^{15}N$-HSQC spectra of uniformly labeled protein. The exceptional purity of the protein should allow its crystallization and detailed structural investigations, and is a prerequisite for its use as a probe in ligand identification and antibody epitope mapping experiments.

Introduction

Natural killer (NK)[1] cells represent a subpopulation of large granular lymphocytes that play a key role in elimination of virally infected and tumor cells [1]. The cytolytic activity of NK cells depends on sophisticated repertoire of activating and inhibitory receptors which are expressed on their surface. NK receptors can be divided into two families by structural homology: the immunoglobulin-like NK receptors (such as KIRs, LILRs and PIRs) and C-type lectin-like NK receptors (such as Ly49s, CD94/NKG2 and NKR-P1) [2,3]. The latter group of receptors appears to be an important component of an alternative missing self-system of NK cells based not on the recognition of MHC class I, but on unusual, apparently carbohydrate-independent lectin–lectin interactions between individual NKR-P1 isoforms and a newly identified family of C-type lectin related (Clr) ligands [4,5]. This receptor system appears to be important for recognition of tumor cells, viral immunity, and immune recognition of cells subjected to genotoxic stress upon administration of chemotherapeutic compounds in rodents [5–7]. In mice, a total of 7 NKR-P1 receptors and 5 Clr ligands have been identified, and shown to be important for both inhibition and activation of certain lymphocyte subsets. In rats, 5 NKR-P1 receptors and 7 Clr ligands have been identified, although the molecular details of interactions and functions of these individual receptor–ligand pairs are mostly still missing [8,9].

The mouse NKR-P1C (mNKR-P1C) belongs to the disulfide-linked homodimeric type II transmembrane C-type lectin-like receptors encoded by the NKR-P1 gene family. This receptor is composed of C-type lectin-like domain connected by a stalk to the transmembrane and cytoplasmic domain [10,11]. The *Nkrp1* (also known as *Klrb1*) gene cluster is located in the NK gene complex (NKC) on chromosome 6 in mice, chromosome 4 in rats, and chromosome 12 in humans [11–13]. The mNKR-P1C receptor corresponds to the NK1.1 alloantigen, and represents one of the most important surface marker of mouse NK cells and NKT cells in C57BL/6 mice detected by monoclonal antibody PK136 [11]. Although mNKR-P1C had been originally cloned as the NK1.1 antigen, analysis revealed that NK1.1 antigen epitope is shared by two receptors, mNKR-P1C and mNKR-P1B [14,15]. The presence of a charged arginine residue in transmembrane segment of mNKR-P1C and absence of cytoplasmatic ITIM motif indicate that

* Corresponding author at: Department of Biochemistry, Faculty of Science, Charles University in Prague, Hlavova 8, CZ-12840 Prague 2, Czech Republic.
E-mail address: bezouska@biomed.cas.cz (K. Bezouška).

[1] *Abbreviations used:* CTLD, C-type lectin-like domain; FT-ICR, Fourier transform ion cyclotron resonance; HSQC, heteronuclear single quantum coherence; IPTG, isopropyl-β-D-thiogalactopyranoside; MALDI, matrix-assisted laser desorption/ionization; MS, mass spectrometry; NK, natural killer; PMSF, phenylmethylsulfonylfluoride; PVDF, polyvinylidenedifluoride; SDS–PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; TFA, trifluoracetic acid; TOF, time of flight.

1046-5928/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.pep.2011.01.013

mNKR-P1C is an activating receptor capable of stimulating the cytolytic process. The activation nature of mNKR-P1C was demonstrated by phosphatidylinositol turnover and Ca^{2+} flux after cross-linking with monoclonal antibody [16], as well as NK cell mediated cytotoxicity and cytokine secretion [17,18]. The signalization pathways of mNKR-P1C remain unknown but it was shown that mNKR-P1C associates with the tyrosine kinase, $p56^{lck}$ through the dicysteine CxCP motif in the cytoplasmatic domain [19]. Although Ocil/Clr-b receptor has been recently identified as the cognate ligand for the inhibitory mNKR-P1B and mNKR-P1D, the physiological ligand of mNKR-P1C remains elusive [20]. However, the molecular nature of strain-dependent NK1.1 alloreactivity in mouse NK cells has been recently clarified in a detailed study performed on BALB/c mice. These mice had a completely functional gene for mNKRP-1C(BA), despite the lack of reactivity with PK136 antibodies [21]. However, this receptor differs in 18 amino acid position in its extracellular part from mNKR-P1C(B6) including the mutation S191T responsible for the loss of PK136 reactivity [21]. Other important mutations involve two structurally critical cysteine residues (C88S and C122S) pointing to principal structural differences between these two receptors [21].

Considering the lack of any knowledge on mNKR-P1C ligands, absence of data on its three dimensional structure, and the need to produce antigen suitable for antibody epitope mapping, it would be desirable to develop a suitable expression system allowing to produce this protein cheaply in sufficient quantities and purity. However, recombinant soluble mNKR-P1C has not been previously produced in any expression system. We now provide a fast and reliable method for expression, refolding and purification of significant quantities of pure, recombinant soluble form of mNKR-P1C(B6). In addition, we describe a detailed optimization of protein production and refolding *in vitro*. The purity of protein preparation was verified by mass spectrometry and its folding state by NMR spectrometry. Unusual stability and high quality of recombinant mNKR-P1C(B6) makes it suitable for studies of its structure and function.

Materials and methods

Materials

Oligonucleotide primers were obtained from Generi Biotech (CZ). Superscript III reverse transcriptase was purchased from Invitrogen (USA), and other enzymes and buffers for DNA manipulations were obtained from New England Biolabs (USA). The pBluescript SK + cloning vector and *Escherichia coli* BL-21 (DE3) Gold strain were from Stratagene (USA), and the pET-30 a(+) expression vector was from Novagene (USA). The Q Sepharose Fast Flow, and Superdex 75 10/300 GL column were from GE Healthcare (Sweden). All other chemicals were purchased form Sigma Aldrich (USA), and were of the highest purity available from this supplier.

Cloning

Total RNA was isolated from spleens of C57BL/6 mice with TRI reagent, and used as a template for single stranded cDNA synthesis using Superscript III reverse transcriptase. DNA fragment coding for mouse NKR-P1C(B6) was amplified by PCR using forward primer 5′-gcc ata tgt cag tta att tag agt gcc ca-3′ containing a *NdeI* site and reverse primer 5′-gca agc tta tca gga gtc att act cgg ggt-3′ containing a *HindIII* site. The PCR product was subcloned into the *SmaI* site of pBluescript SK + cloning vector. The resulting plasmid was digested with *NdeI* and *HindIII* and the fragment containing mNKR-P1C(B6) was ligated into the pET-30 a(+) expression vector linearized with the same restriction endonucleases. The correctness of the sequence was confirmed by DNA sequencing. The obtained construct pET-30-NKRP1C(B6) coded for recombinant protein S89 to S223 of mouse NKR-P1C(B6) with a calculated molecular mass of 15,409 Da.

Protein expression

For the production of protein, construct pET-30-NKRP1C(B6) was transformed into *E. coli* BL-21 (DE3) Gold strain. A starter culture for producing mNKR-P1C(B6) was established overnight in a 5 mL of LB medium with 50 μg/mL kanamycin and 12.5 μg/mL tetracycline. Large-scale protein production was carried out by adding 5 mL of the starter culture to 500 mL of kanamycin and tetracycline supplemented LB medium in a 2 L Erlenmeyer flask. Cells were grown at 37 °C with shaking at 220 rpm. When the cell density reached OD_{600} of 0.8, expression of the recombinant protein was induced by addition of isopropyl-β-D-thiogalactopyranoside to a final concentration of 0.01 mM. After induction, cells were grown for 4 h and then harvested by centrifugation at 6000*g* for 10 min. Large-scale protein production of ^{15}N-labeled mNKR-P1C(B6) was

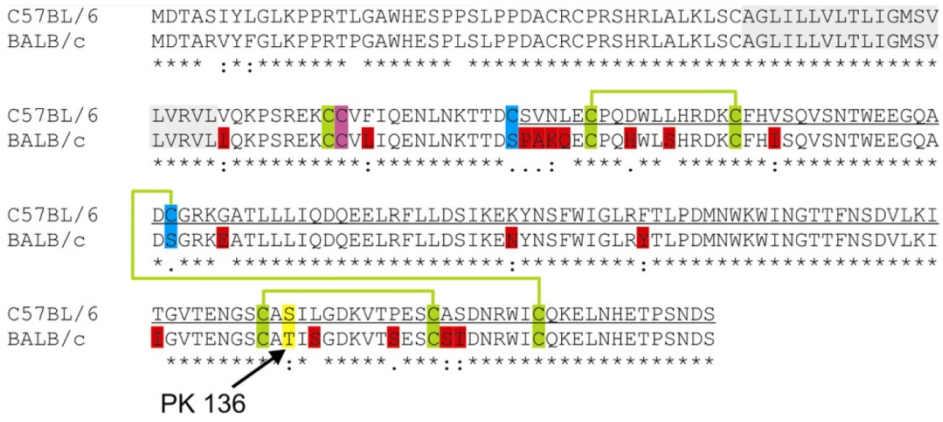


Fig. 1. Sequence alignment of mouse NKR-P1C(B6) and NKR-P1C(BA) from the two respective mouse strains, C57BL/6 and BALB/c. Individual amino acid differences within the entire extracellular domains of the two receptors are marked in red, the conserved cysteines characteristic for the CTLD are indicated in green, the two cysteines with mutations in blue, the conserved cysteines involved in interchain dimerization in magenta, and the transmembrane segment in gray. Mutation S191T responsible for the loss if immunoreactivity with monoclonal antibody PK136 characterizing NK1.1 antigen [21] is shown in yellow. Soluble receptor described here contained approximately half of the membrane-connecting stalk and entire CTLD encompassing amino acids S89–S223 of the C57BL/6 sequence (underlined). The putative disulfide arrangement in CTLD is marked by linking the corresponding cystines.

carried in M9 minimal medium [20] containing antibiotics as described above, and $^{15}NH_4Cl$ (Cambridge Isotope Laboratories) as a sole nitrogen source.

Purification of mNKR-P1C(B6) from inclusion bodies and refolding

The cell pellet from 1 L of bacterial culture (2.5 g of wet weight cells) was resuspended in 20 mL of buffer containing 50 mM Tris–HCl (pH 7.4), 25% sucrose (w/v), 1 mM EDTA, 1 mM NaN_3, 1 mM phenylmethylsulfonylfluoride (PMSF), 1 μM leupeptin and 1 μM pepstatin. The cell suspension was subjected to 4 cycles of freeze/thaw at −80 and 37 °C. The cells were lysed by sonication on ice for 2 min, and the disrupted cells were then supplemented with 100 μg of DNase I, 100 μg RNase I and 400 μL of 1 M $MgCl_2$, and incubated at room temperature for 20 min. Following centrifugation at 15,000*g* for 10 min, the pellet was washed with 20 mL of buffer containing 50 mM Tris–HCl (pH 7.4), 0.5% Triton X-100 (v/v), 100 mM NaCl, 1 mM 2-mercaptoethanol, 1 mM NaN_3, 1 mM PMSF, 1 μM leupeptin and 1 μM pepstatin. This step was repeated once with 10 mL of the latter buffer without Triton X-100. Following the final centrifugation at 15,000*g* for 10 min, the washed inclusion bodies were solubilized in 4 mL of 6 M guanidine–HCl (pH 8.0) containing 10 mM DTT, 1 μM leupeptin and 1 μM pepstatin by incubation at 40 °C for 1 h. The solution was cleared from insoluble material by centrifugation at 50,000*g* for 30 min. The mNKR-P1C(B6) protein was refolded by rapid dilution into a 100-fold excess (400 ml) of refolding buffer containing 50 mM Tris–HCl (pH 8.5), 0.4 M L-arginine, 100 mM $CaCl_2$, 9 mM cysteamine, 3 mM cystamine, 1 mM NaN_3 and 1 mM PMSF. The refolding mixture was then dialyzed twice at 4 °C against 8 L of 15 mM Tris–HCl (pH 8.5), 9 mM NaCl and 1 mM NaN_3 for 6 h.

Purification of recombinant mNKR-P1C(B6)

The refolded protein was loaded onto a Q-Sepharose FF column (1.6 × 14.0 cm, GE Healthcare) equilibrated in the dialysis buffer, and eluted by linear gradient of NaCl from 9 mM to 1 M. Fractions were analyzed by SDS–PAGE under reducing and nonreducing conditions, and those containing folded protein were pooled and concentrated by ultrafiltration using cellulose membranes with a 10 kDa cut-off (Millipore). The concentrated protein was finally purified by gel filtration on a Superdex 75 10/300 GL column (GE Healthcare), and dialyzed against 15 mM Tris–HCl (pH 8.0), 150 mM NaCl and 1 mM NaN_3. Protein concentration was measured using the Bradford assay (Bio-Rad) with bovine serum albumin as a standard [22].

Analytical ultracentrifugation

Sedimentation analysis was performed using a ProteomeLabXL-I analytical ultracentrifuge equipped with An50Ti rotor (Beckman-Coulter). The mNKR-P1C(B6) protein (0.4 mg/ml) was dialyzed against 10 mM HEPES (pH 7.0), 50 mM NaCl and 1 mM NaN_3 buffer, used also as a reference. Sedimentation velocity experiment was carried out at 40,000 rpm and 20 °C, absorbance scans were recorded at 300 nm in 10 min intervals with 30 μm spatial resolution. Buffer density and mNKR-P1C(B6) partial specific volume were estimated in SEDNTERP 1.09 (www.jphilo.mailway.com). Data were analyzed with SEDFIT 12.1 [23]. The sedimentation equilibrium experiment was performed with mNKR-P1C(B6) concentration 0.1 mg/ml at 27–30–33,000 rpm at 20 °C. Absorbance data were collected at 295 nm by averaging 20 scans with 10 μm spatial resolution after 16 h of achieving equilibrium and were globally analyzed with SEDPHAT 8.2 using a non-interacting discrete species model [24].

NMR measurements

All NMR experiments were recorded at 25 °C on Bruker Avance III 600 MHz spectrometer equipped with a TCI cryoprobe (Bruker Daltonics). ^1H–^{15}N-HSQC spectra were used as a routine check of protein folding and stability. NMR experiments were carried out using uniformly labeled ^{15}N protein sample of 0.3 mM concentration in buffer containing 15 mM PIPES (pH 6.8), 50 mM NaCl, 1 mM NaN_3 and 10% D_2O. NMR data were processed and analyzed using the NMRPipe and Sparky software [25,26].

Mass spectrometric analysis and N-terminal sequencing

Accurate measurement of intact protein mass was performed on a 9.4 T Apex-ULTRA Fourier transformed mass spectrometer (Bruker Daltonics). Protein was first desalted on a Microtrap column (Michrome Bioresources) before electrospray ionization.

For sequence mapping of tryptic digests with MALDI TOF MS, Coomassie blue-stained protein bands were excised from the SDS–PAGE gel, cut into small pieces, and washed several times with 0.1 M 4-ethylmorpholine acetate buffer (pH 8.3) containing 50% acetonitrile. After complete destaining, the protein was reduced by 50 mM tris[(2-carboxyethyl)phosphine hydrochloride] (TCEP) at 90 °C for 5 min and alkylated by treatment with 50 mM iodacetamide for 20 min in the dark. The gel pieces were further washed with acetonitrile and then rehydrated with water. This step was repeated three times. The gel pieces were partly dried in vacuum and then rehydrated in a cleavage buffer containing 50 mM 4-ethylmorpholine acetate, 10% acetonitrile and 1 μl of

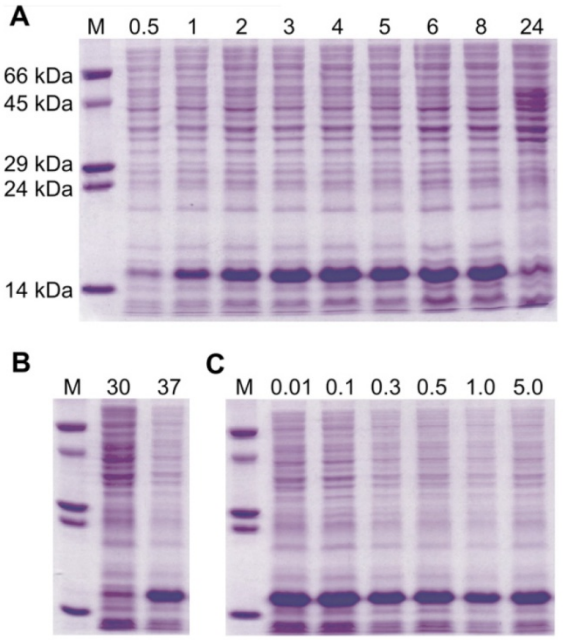


Fig. 2. Optimization of protein production using SDS–PAGE analysis of cell lysates. (A) In the optimization of the duration of protein production, aliquots of cell culture were harvested at 0.5, 1, 2, 3, 4, 5, 6, 8 and 24 h post induction and equal volumes were loaded for SDS–PAGE analysis. The recombinant mNKR-P1C(B6) was identified as band with an apparent molecular weight of approximately 16 kDa and protein production reached sufficient amount about 4 h after protein induction with IPTG. (B) Optimization of production temperature showed that mNKR-P1C(B6) was poorly produced at 30 °C whereas good yield was achieved at 37 °C. (C) Optimization of IPTG concentration, the numbers over the lines represent the final concentration of IPTG.

trypsin (5 ng/μl). Positive-ion mass spectra were measured on a Bruker ULTRAFLEX III TOF mass spectrometer (Bruker Daltonics). The spectra of the peptides obtained after proteolytic digestion were measured in reflectron mode by using α-cyano-4-hydroxycinnamic acid in aqueous 40% acetonitrile-0.2% TFA (10 mg/ml) as a MALDI matrix.

For N-terminal sequencing, mNKR-P1C(B6) was electroblotted from 15% SDS–PAGE gel onto a PVDF membrane and analyzed by automated cycles of Edman degradation in a Procise 491 Protein Sequencer (Applied Biosystems).

Results

Optimization of protein expression in E. coli

Recombinant soluble mNKR-P1C(B6) protein has not been previously produced in any expression system. Here, we describe a fast and efficient method for the expression, refolding and purification of the soluble form of mNKR-P1C(B6) containing the part of stalk region and the extracellular ligand binding domain encompassing amino acids S89–S223 (Fig. 1). To express the recombinant protein, total RNA was isolated from spleen of C57BL/6 mouse, and purified RNA was used for first-strand cDNA synthesis by RT-PCR. DNA fragment of mNKR-P1C(B6) was amplified by PCR using cDNA as template and subcloned into pET-30 a(+) expression vector. The sequence was verified by DNA sequencing. The construct was then transformed into *E. coli* strain BL(21)DE3 Gold. For high-level expression, the time of production, IPTG concentration and induction temperature were optimized by SDS–PAGE of cell lysates (Fig. 2). The recombinant mNKR-P1C(B6) was identified as band with an apparent molecular weight of approximately 16 kDa by sequence mapping of tryptic digests with MALDI TOF MS. The optimization of the duration of protein production (Fig. 2A) showed that recombinant mNKR-P1C(B6) was rapidly expressed following 30 min and reached sufficient amount about 4 h after protein induction with IPTG. The maximum yield of mNKR-P1C(B6) was achieved at production temperature of 37 °C (Fig. 2B) and optimal IPTG concentration was rather low, 0.01 mM (Fig. 2C). However, essentially no protein was produced in the absence of IPTG (not shown). Therefore, the large-scale production of recombinant mNKR-P1C(B6) was performed under these conditions. As expected, recombinant protein was deposited in inclusion bodies, and correctly folded protein had to be prepared by *in vitro* refolding.

Protein refolding

The method used for isolation of inclusion bodies involved resuspension of the bacterial pellet in cell lysis buffer containing 25% sucrose followed by centrifugation and two further washes of cell debris in buffer containing 0.5% Triton X-100, and a final wash with the same buffer without detergent. The washed inclusion bodies were solubilized in 6 M guanidine–HCl (pH 8.0) containing 10 mM DTT, and mNKR-P1C(B6) was refolded by rapid

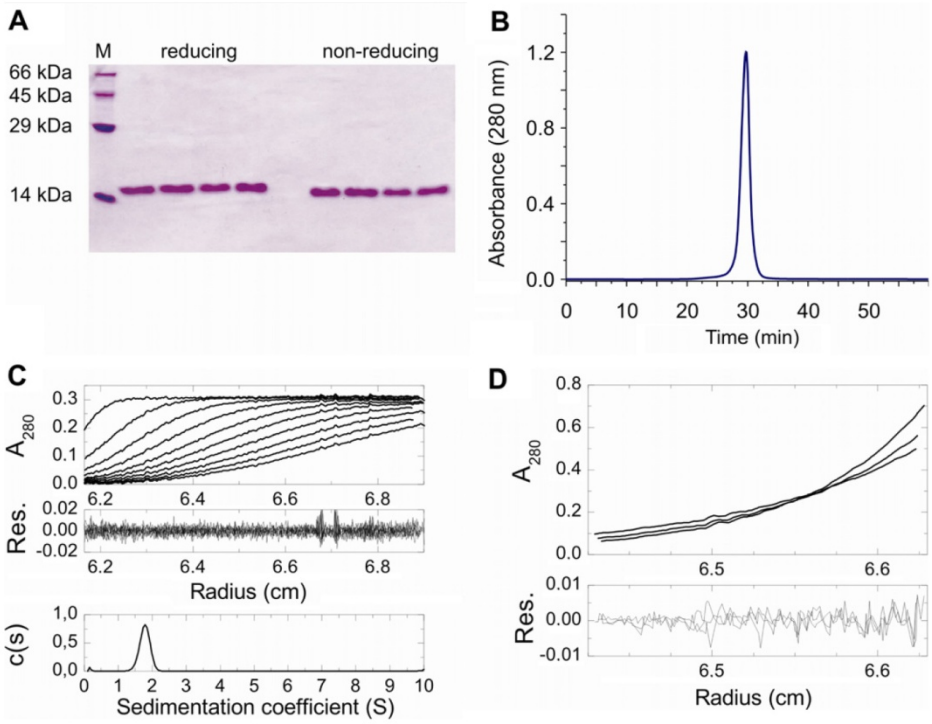


Fig. 3. Characterization of the soluble form of mNKR-P1C(B6). (A) SDS–PAGE analysis of the fractions eluted during the purification by one-step anion exchange chromatography on Q-Sepharose FF. The notable shift in mobility under nonreducing conditions was observed due to compact arrangement of the protein cross-linked by three disulfide bridges. (B) Elution profile of refolded mNKR-P1C(B6) from size-exclusion chromatography. The protein was injected on Superdex 75 10/300 GL column run at 0.4 ml/min in 15 mM Tris–HCl (pH 8.0), 150 mM NaCl and 1 mM NaN_3. (C) Size distribution of sedimenting mNKR-P1C(B6) species. Sedimentation velocity experiment was carried out at 40,000 rpm at 20 °C and absorbance scans were recorded every 10 min. Fitted data (upper panel) with residual plot (middle panel) showing goodness of fit are shown together with calculated continuous size distribution c(s) of sedimenting species (lower panel). (D) Equilibrium sedimentation distribution of mNKR-P1C(B6). Protein sample (0.1 mg/ml) was spun at 20 °C at 27–30–33,000 rpm and its sedimentation was monitored after 16 h with absorbance optics. Upper panel shows absorbance data (circles) with fitted curves (non-interacting discrete species model, lines), lower panel shows residuals derived from the fitted data.

dilution. Many different refolding conditions were investigated (data not shown) differing in concentration of L-arginine, pH, and redox shuffling agents as well as their ratios. The highest refolding yield was obtained using 100-fold dilution into refolding buffer containing 50 mM Tris–HCl (pH 8.5), 0.4 M L-arginine, 100 mM $CaCl_2$, 9 mM cysteamine, 3 mM cystamine, 1 mM NaN_3 and 1 mM phenylmethanesulfonyl fluoride, followed by two rounds of dialysis against 8 L of 15 mM Tris–HCl (pH 8.5), 9 mM NaCl and 1 mM NaN_3 for 6 h at 4 °C.

Purification of mNKR-P1C(B6) to homogeneity

The refolded mNKR-P1C(B6) purifies using chromatography on a single anion-exchange column. Protein prepared by this protocol appeared homogenous by SDS–PAGE under reducing and nonreducing conditions (Fig. 3A), and notable shift in mobility under nonreducing conditions was observed due to compact arrangement of the protein cross-linked by three disulfide bridges. The final yield was estimated as 20 mg of pure recombinant mNKR-P1C(B6) protein per liter of bacterial culture, and its purity reached more than 98%. The yields and purities of target protein during procedures of refolding and purification are summarized in Table 1.

Size-exclusion chromatography with a Superdex 75 10/300 GL column used for final purification of the monodisperse protein suggested that mNKR-P1C(B6) elutes exclusively as a monomer (Fig. 3B). This finding was further corroborated by sedimentation analysis in an analytical ultracentrifuge. Sedimentation coefficient of mNKR-P1C(B6) corrected to standard conditions (water, 20 °C, zero protein concentration) has value of 1.82 S, as determined by sedimentation velocity experiment (Fig. 3C). This value corresponds well to the mass of a monomer of moderately elongated shape, however, it would be too much low for the mass of any dimer. Finally, sedimentation equilibrium experiment was performed (Fig. 3D) that allows a direct calculation of molecular mass without the need for calibration or interaction with any matrix, yielded the weight average molecular mass of 15,047 ± 121 Da for single ideal species model, correlating well with the expected mass of mNKR-P1C(B6) monomer (15,409 Da).

Analysis of the covalent structure of recombinant mNKR-P1C(B6)

N-terminal sequencing revealed that the protein was produced in native form devoid of the initiation methionine. Otherwise, 16 cycles of automated Edman degradation provided a sequence SVNLEXPQDWLLHRDK, which was in a complete agreement with the expected amino acid sequence except for a cysteine residue at position 6. This amino acid residue, when not chemically modified, is destroyed under the conditions of Edman chemistry, and therefore could not be seen on the recorded sequence. Monoisotopic mass of intact protein mNKR-P1C(B6) $[M+H]^+$ determined by high resolution FT-ICR MS was 15,394.38 Da (Fig. 4A). This is in an excellent agreement with the predicted monoisotopic mass of the protein mNKR-P1C(B6) $[M+H]^+$ 15394.18 Da calculated with an assumption that all three of the expected intramolecular disulfide bonds were closed, and that the entire polypeptide was devoid of the N-terminal initiation methionine as proved above by N-terminal sequencing results (cf. above).

Table 1
Course of the purification followed by SDS–PAGE. All protein amounts relate to 1 L of bacterial culture.

Purification step	Total protein (mg)	Target protein (mg)	Purity (%)	Yield (%)
Inclusion bodies isolation*	300	210	70	100
Protein refolding	94	80	85	38
Ion-exchange chromatography	35	33	94	16
Size-exclusion chromatography	20	20	<98	10

* From 1 L of bacterial culture 2.5 g of wet weight cells was obtained.

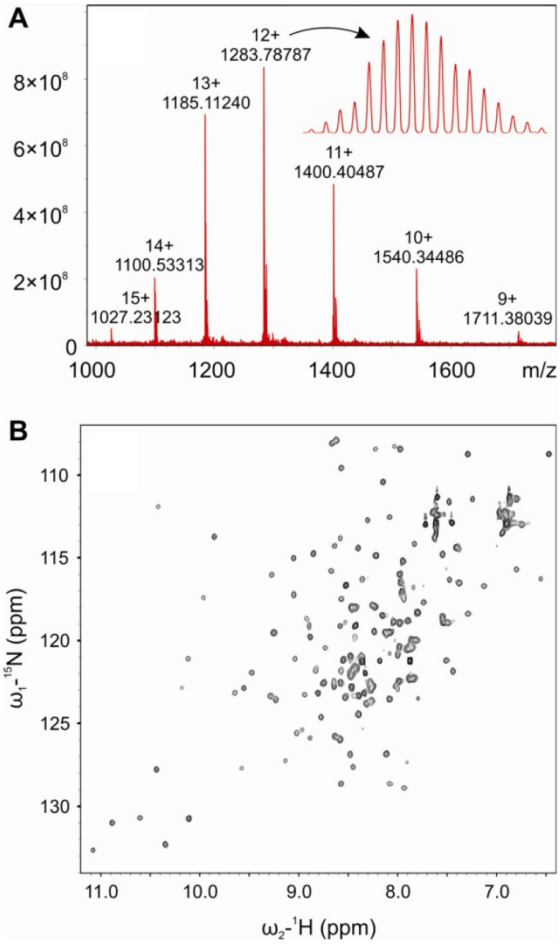


Fig. 4. High resolution ESI/FT-ICR MS and NMR analysis of recombinant mNKR-P1C(B6). (A) Full mass spectrum of multiply charged mNKR-P1C(B6). The charge-state envelope was deconvoluted to determine the mass of protein which corresponds to the expected mass with all cysteine forming disulfide bridges. (B) ^1H–^{15}N-HSQC spectrum of 0.3 mM solution of the uniformly ^{15}N-labeled mNKR-P1C(B6) indicates good dispersion of the backbone and side-chain signals. The spectrum was acquired at 600 MHz at 25 °C. The buffer conditions are 10 mM PIPES (pH 6.8), 50 mM NaCl and 1 mM NaN_3.

Evidence for proper protein folding

A ^1H–^{15}N-HSQC spectrum was recorded on ^{15}N uniformly labeled mNKR-P1C(B6) in buffer containing 10 mM PIPES (pH 6.8), 50 mM NaCl and 1 mM NaN_3. The ^1H–^{15}N-HSQC of the 0.3 mM solution of this protein (Fig. 4B) indicates good dispersion of the backbone and side chain signals with excellent 5 ppm chemical shift dispersion in proton dimension. The N–H indole proton of the six tryptophan residues of mNKR-P1C(B6) are situated in the lower left corner of the spectrum, and NH_2 signals belonging to asparagine and glutamine residues are in the upper right region

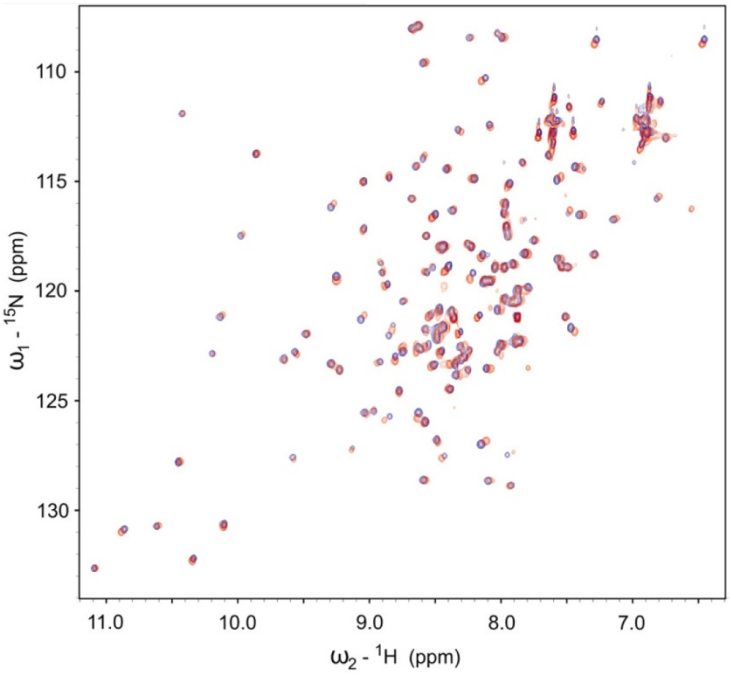


Fig. 5. Structural stability of recombinant mNKR-P1C(B6) monitored by NMR. Two different $^{1}H–^{15}N$-HSQC spectra of mNKR-P1C(B6) were acquired before (blue) and after (red) 90 days of incubation at 4 °C. No significant change in the spectra indicates good stability of protein preparation.

of the spectrum. When the same sample was analyzed after 90 days of storage at 4 °C, essentially identical spectrum was obtained (Fig. 5), which indicates high stability of the protein preparation. The $^{1}H–^{15}N$-HSQC spectra of several different batches were almost identical, indicating reproducibility of the refolding and purification protocol (data not shown). This demonstrates that purified recombinant mNKR-P1C(B6) is a well folded protein amenable to structure determination by NMR.

Discussion

Recently, there has been a renewed interest in the NKR-P1 proteins, a group of C-type lectin-like receptors of natural killer cells, that have been shown to represent an alternative "missing-self" recognition system [4]. In particular, mouse NKR-P1C receptor appears to be interesting beyond the basic biology of natural killer cells. The mouse NKR-P1C receptor is by far the best known antigen within this family since it has been used as one of the most important surface marker of mouse NK cells for more than three decades [17,21]. The strain-specific expression of this surface marker has been a mystery for a long time until its molecular analysis revealed it to be actually an artefact related to the fine specificity of the detecting monoclonal antibody PK-136 [4,21]. This clarification has sparked efforts into production of alternative monoclonal antibodies against the common epitopes of this antigen that has been proven to be functionally important for activation of mouse NK cells [16,18]. Thus, the readily available soluble mouse NKR-P1C(B6) protein would be useful for testing of the binding specificity of individual antibody clones. Since the physiological ligand for mNKR-P1C(B6) has not been described, such protein would also be invaluable as a probe in ligand identification experiments. Furthermore, despite many years of research and availability of molecular models [27], the exact three-dimensional structure of any of the NKR-P1 antigens has not been published, and thus large amounts of soluble mNKR-P1C(B6) protein would be required in order to initiate crystallization trials and structural NMR experiments. However, no report has been published so far describing the recombinant expression of this particular antigen.

Here we describe a simple and efficient recombinant expression system allowing the production and refolding of an extended extracellular portion of mouse NKR-P1C(B6) using simple and fast protocol. After induction of bacterial expression, large amount of the desired protein was produced and precipitated into inclusion bodies that allowed us to recover the intact protein in good yield and relatively high purity. The precipitated protein was solubilized in denaturing conditions, and refolded using a rapid dilution method. The refolding mixture contained L-arginine as "low molecular weight chaperone" additive, and a cysteamine/cystamine redox shuffling pair, both shown previously to be useful for *in vitro* refolding of related C-type lectin-like receptors such as CD69 [28]. Successful refolding of NKR-P1C(B6) could be achieved under a large number of experimental conditions, of which the most suitable were selected on the basis of the lowest production costs. In contrast to a protocol reported previously for rat NKR-P1A [29], our method utilize an expression construct devoid of any targeting signals or purification tags, and yet yields highly pure protein in just two chromatography steps.

Natural variants of mouse NKR-P1C receptors belong to a rather difficult proteins from the point of view of their biosynthesis. Transcription of the corresponding gene is limited to natural killer cells and a few additional lymphocyte subsets, and is tightly regulated, as would be expected for surface receptor capable of delivering important decisions for the cell that expresses it. Complex regulatory mechanisms involved in the lineage-specific control of NKR-P1C expression in NK lymphocytes have been described [30]. Moreover, in case of mNKR-P1C(B6) at least three transcriptional variants have been detected: the full length isoform 1, isoform 2 shortened by three amino acids (including deletion of one of the cysteine residues in the receptor extracellular stalk region), and isoform 3 shortened by additional three amino acids. The existence

of multiple (and variable) amount of cysteines in the extracellular portion of the receptor makes it very difficult to select an appropriate fragment for recombinant expression, as well as the suitable expression system. Furthermore, a complex system of posttranslational modifications occurs in the receptor molecule involving not only disulfide pairing, but also glycosylation and phosphorylation. Despite all these hurdles, we demonstrate here that it is possible to prepare milligram amounts of properly folded and highly stable soluble extracellular part of this receptor using rapid and economic bacterial expression system. The produced protein has been found to be monomeric what somewhat limits its use in ligand identification studies compared to the natural, dimeric form of the receptor. In our current efforts we are trying to make use of the optimized expression system described above in order to produce the entire extracellular portion of mouse NKR-P1C(B6) and related proteins in a dimeric form as well.

Conclusion

In this paper, we have described a method for producing milligram quantities of recombinant mouse NKR-P1C(B6). This is the first report of a recombinant expression of a soluble form of this important receptor. The monomeric mNKR-P1C(B6) can be easily expressed and purified in high yield (20 mg per liter bacterial culture) over a period of 4–6 days using common biochemical laboratory equipment. The SDS–PAGE and size-exclusion chromatography demonstrate the monodispersity of the protein preparation, and NMR data indicates that purified mNKR-P1C(B6) is well folded with a remarkable stability *in vitro*. The recombinant mNKR-P1C(B6) prepared using our protocol should be suitable for structure determination by NMR, crystal structure analysis or for identification of the physiological ligand(s). Moreover, highly efficient production and refolding protocols that we have developed open the way for the production of mNKR-P1C(B6) with the extended neck region under similar conditions. Finally, the parallel production of mNKR-P1C(BA) from BALB/c mice and its structural investigation (to be reported separately) may shed light on the important structural differences between these two functionally closely related receptors that diverged significantly during the recent evolution.

Author contributions

D.R., O.V., and K.B. designed research; D.R., D.K. and J.C. performed research; D.R., J.C., P.N., and K.B. analyzed data, D.R., O.V., and K.B. wrote the paper.

Acknowledgments

This research was supported by Czech Science Foundation (Grants 303/09/0477, 305/09/H008, and P207/10/1040), the Grant Agency of Academy of Sciences of the Czech Republic (IAA500200620), and by Ministry of Education of the Czech Republic (MSM_21620808, LC7017, LC545, 1M0505).

References

[1] E. Vivier, E. Tomasello, M. Baratin, T. Walzer, S. Ugolini, Functions of natural killer cells, Nat. Immunol. 9 (2008) 503–510.
[2] L.L. Lanier, NK cell recognition, Annu. Rev. Immunol. 23 (2005) 225–274.
[3] K. Natarajan, N. Dimasi, J. Wang, R.A. Mariuzza, D.H. Margulies, Structure and function of natural killer cell receptors: multiple molecular solution to self, nonself discrimination, Annu. Rev. Immunol. 20 (2002) 853–885.
[4] A. Mesci, B. Ljutic, A.P. Makrigiannis, J.R. Carlyle, NKR-P1 biology. From prototype to missing self, Immunol. Res. 35 (2006) 13–26.
[5] K. Iizuka, O. Naidenko, B. Plougastel, D. Fremont, W.M. Yokoyama, Genetically-linked C-type lectin-related ligands for the NKR-P1 NK cell receptors, Nat. Immunol. 4 (2003) 801–807.
[6] S. Voigt, A. Mesci, J. Ettinger, J.H. Fine, P. Chen, W. Chou, J.R. Carlyle, Cytomegalovirus evasion of innate immunity by subversion of the NKR-P1B: Clr-b missing self axis, Immunity 26 (2007) 617–627.
[7] J.H. Fine, P. Chen, A. Mesci, D.S.J. Allan, S. Gasser, D.H. Raulet, J.R. Carlyle, Chemotherapy-induced genotoxic stress promotes sensitivity to natural killer cell cytotoxicity by enabling missing-self recognition, Cancer Res. 70 (2010) 7102–7113.
[8] B. Plougastel, K. Matsumoto, C. Dubbelde, W.M. Yokoyama, Cloning of Clr, a new family of lectin-like genes localized between mouse Nkrp1a and Cd69, Immunogenetics 53 (2001) 592–598.
[9] L.M. Flornes, G. Nylenna, P.C. Saether, M.R. Daws, E. Dissen, S. Fossum, The complete inventory of receptors encoded by the rat natural killer cell gene complex, Immunogenetics 62 (2010) 521–530.
[10] R. Giorda, M. Trucco, Mouse NKR-P1. A family of genes selectively coexpressed in adherent lymphokine-activated killer cells, J. Immunol. 147 (1991) 1701–1708.
[11] J.C. Ryan, J. Turck, E.C. Niemi, W.M. Yokoyama, W.E. Seaman, Molecular cloning of the NK1.1 antigen, a member of the NKR-P1 family of natural killer cell activation molecules, J. Immunol. 149 (1992) 1631–1635.
[12] J.C. Ryan, W.E. Seaman, Divergent functions of lectin-like receptors on NK cells, Immunol. Rev. 155 (1997) 79–89.
[13] W.M. Yokoyama, J.C. Ryan, J.J. Hunter, H.R. Smith, M. Stark, W.E. Seaman, CDNA cloning of mouse NKR-P1 and genetic linkage with LY-49. Identification of a natural killer cell gene complex on mouse chromosome 6, J. Immunol. 147 (1991) 3229–3236.
[14] J.R. Carlyle, A. Martin, A. Mehra, L. Attisano, F.W. Tsui, J.C. Zúñiga-Pflücker, Mouse NKR-P1B, a novel NK1.1 antigen with inhibitory function, J. Immunol. 162 (1999) 5917–5923.
[15] S.K. Kung, R.C. Su, J. Shannon, R.G. Miller, The NKR-P1B gene product is an inhibitory receptor on SJL/J NK cells, J. Immunol. 162 (1999) 5876–5887.
[16] J.C. Ryan, E.C. Niemi, R.D. Goldfien, J.C. Hiserodt, W.E. Seaman, NKR-P1, an activating molecule on rat natural killer cells, stimulates phosphoinositide turnover and a rise in intracellular calcium, J. Immunol. 147 (1991) 3244–3250.
[17] F.M. Karlhofer, W.M. Yokoyama, Stimulation of murine natural killer (NK) cells by a monoclonal antibody specific for the NK1.1 antigen. IL-2-activated NK cells possess additional specific stimulation pathways, J. Immunol. 146 (1991) 3662–3673.
[18] H. Arase, N. Arase, T. Saito, Interferon gamma production by natural killer (NK) cells and NK1.1+ T cells upon NKR-P1 cross-linking, J. Exp. Med. 183 (1996) 2391–2396.
[19] B. Ljutic, J.R. Carlyle, D. Filipp, R. Nakagawa, M. Julius, J.C. Zúñiga-Pflücker, Functional requirements for signaling through the stimulatory and inhibitory mouse NKR-P1 (CD161) NK cell receptors, J. Immunol. 174 (2005) 4789–4796.
[20] J.R. Carlyle, A.M. Jamieson, S. Gasser, C.S. Clingan, H. Arase, D.H. Raulet, Missing self-recognition of Ocil/Clr-b by inhibitory NKR-P1 natural killer cell receptors, Proc. Natl. Acad. Sci. USA 101 (2004) 3527–3532.
[21] J.R. Carlyle, A. Mesci, B. Ljutic, S. Belanger, L.-H. Tai, E. Rouselle, A.D. Troke, M.-F. Proteau, A.P. Makrigiannis, Molecular and genetic basis for strain-dependent NK1.1 alloreactivity of mouse NK cells, J. Immunol. 176 (2006) 7511–7524.
[22] M.M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, Anal. Biochem. 72 (1976) 248–254.
[23] P. Schuck, Size distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modelling, Biophys. J. 78 (2000) 1606–1619.
[24] P. Schuck, On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation, Anal. Biochem. 320 (2003) 104–124.
[25] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, A. Bax, NMRPipe: a multidimensional spectral processing system based on UNIX pipes, J. Biomol. NMR 6 (1995) 277–293.
[26] T.D. Goddard, D.G. Kneller, SPARKY3, University of California, San Francisco.
[27] Ž. Sovová, V. Kopecký Jr., T. Pazderka, K. Hofbauerová, D. Rozbeský, O. Vaněk, K. Bezouška, R. Ettrich, Structural analysis of natural killer cell receptor protein 1 (NKR-P1) extracellular domains suggests a conserved long loop region involved in ligand specificity, J. Mol. Model. (2010) doi:10.1007/s00894-010-0837-y.
[28] O. Vaněk, M. Nálezková, D. Kavan, I. Borovičková, P. Pompach, P. Novák, V. Kumar, L. Vannucci, J. Hudeček, K. Hofbauerová, V. Kopecký, J. Brynda, P. Kolenko, J. Dohnálek, P. Kaderávek, J. Chmelík, L. Gorčík, L. Žídek, V. Sklenář, K. Bezouška, Soluble recombinant CD69 receptor optimized to have an exceptional physical and chemical stability display prolonged circulation and remain intact in the blood of mice, FEBS J. 275 (2008) 5589–5606.
[29] H. Kogelberg, A.M. Lawson, F.W. Muskett, R.A. Carruthers, T. Feizi, Expression in *Escherichia coli*, folding in vitro, and characterization of the carbohydrate recognition domain of the natural killer cell receptor NKR-P1A, Protein Expr. Purif. 20 (2000) 10–20.
[30] B. Ljutic, J.R. Carlyle, J.C. Zuniga-Pflucker, Identification of upstream *cis*-acting regulatory elements controlling lineage-specific expression of the mouse NK cell activation receptor, NKR-P1C, J. Biol. Chem. 278 (2003) 31909–31917.

PAPER II

Sovova, Z., Kopecky, V., Jr., Pazderka, T., Hofbauerova, K., Rozbesky, D., Vanek, O., Bezouska, K., Ettrich, R

Structural analysis of natural killer cell receptor protein 1 (NKR-P1) extracellular domains suggests a conserved long loop region involved in ligand specificity.

J. Mol. Model. **17**, 1353-70 (2011)

My contribution to the publication: *research performing (cloning, protein expression, refolding in vitro, purification, protein charecterization), data collection, data analysis and interpretation, manuscript writing*

J Mol Model
DOI 10.1007/s00894-010-0837-y

ORIGINAL PAPER

Structural analysis of natural killer cell receptor protein 1 (NKR-P1) extracellular domains suggests a conserved long loop region involved in ligand specificity

Žofie Sovová · Vladimír Kopecký Jr. · Tomáš Pazderka · Kateřina Hofbauerová · Daniel Rozbeský · Ondřej Vaněk · Karel Bezouška · Rüdiger Ettrich

Received: 18 July 2010 / Accepted: 24 August 2010

Abstract Receptor proteins at the cell surface regulate the ability of natural killer cells to recognize and kill a variety of aberrant target cells. The structural features determining the function of natural killer receptor proteins 1 (NKR-P1s) are largely unknown. In the present work, refined homology models are generated for the C-type lectin-like extracellular domains of rat NKR-P1A and NKR-P1B, mouse NKR-P1A, NKR-P1C, NKR-P1F, and NKR-P1G, and human NKR-P1 receptors. Experimental data on secondary structure, tertiary interactions, and thermal transitions are acquired for four of the proteins using Raman and infrared spectroscopy. The experimental and modeling results are in agreement with respect to the overall structures of the NKR-P1 receptor domains, while suggesting functionally significant local differences among species and isoforms. Two sequence regions that are conserved in all analyzed NKR-P1 receptors do not correspond to conserved structural elements as might be expected, but are represented by loop regions, one of which is arranged differently in the constructed models. This region displays high flexibility but is anchored by conserved sequences, suggesting that its position relative to the rest of the domain might be variable. This loop may contribute to ligand-binding specificity via a coupled conformational transition.

Keywords Cladogram · Differential scanning calorimetry · FTIR · Molecular dynamics · Raman spectroscopy · RMSF · Thermal dynamics · Topology · Two-dimensional correlation analysis

Ž. Sovová · R. Ettrich (✉)
Laboratory of Structural Biology, Institute of Systems Biology and Ecology, Academy of Sciences of the Czech Republic,
Zámek 136,
37333 Nové Hrady, Czech Republic
e-mail: ettrich@nh.usbe.cas.cz

Ž. Sovová · R. Ettrich
Faculty of Sciences, University of South Bohemia,
Zámek 136,
37333 Nové Hrady, Czech Republic

V. Kopecký Jr. · T. Pazderka · K. Hofbauerová
Institute of Physics, Faculty of Mathematics and Physics,
Charles University in Prague,
Ke Karlovu 5,
12116 Prague 2, Czech Republic

K. Hofbauerová · D. Rozbeský · O. Vaněk · K. Bezouška
Institute of Microbiology,
Academy of Sciences of the Czech Republic,
Vídeňská 1083,
14220 Prague 4, Czech Republic

D. Rozbeský · O. Vaněk · K. Bezouška
Department of Biochemistry, Faculty of Science,
Charles University in Prague,
Albertov 2030,
12840 Prague 2, Czech Republic

Published online: 14 September 2010

Introduction

Natural killer (NK) cells are large granular lymphocytes able to recognize and kill a large variety of target cells and to regulate the reactions at the interface of innate and adaptive immunity through secretion of lymphokines and by direct killing [1, 2]. The cytotoxic activity of NK cells is tightly regulated *via* activating and inhibiting cell-surface receptors, one group of them being the NK cell lectin-like receptor proteins, one subtype of which are the NKR-P1s. The NKR-P1s are transmembrane glycoproteins classified as type II due to their external C-terminus, with an extracellular C-type lectin-like domain (CTLD) and a short cytoplasmic domain. To initiate the NK cell response, activating NK cell receptors recognize a diverse range of

ligands including cytokines, antibody Fc domains and other proteins and saccharides presented by target cells. Inhibitory NK cell receptors recognize MHC class I molecules and other proteins serving as a marker of cell health. Tumor, virally infected, stressed or otherwise damaged cells that can resist T cell mediated immunity because of the low levels of these markers expressed at their surface are no longer protected by their inhibitory signals and this may lead to NK cell activation and an elimination of the target cells [3]. Functions of NKR-P1s were enigmatic until their ligands were found to be closely related family of Clr lectin-like receptors [2]. Despite NKR-P1 family include both activating and inhibitory receptors, the CTLDs of NKR-P1 family members share considerable homology; thus the structural origins of ligand-binding specificity are of interest. In the present work, multiple sequence alignment, protein homology modeling, and molecular dynamics simulations are combined with protein expression, vibrational spectroscopy, and thermal analysis to examine evolutionary and structural divergences within the family of NKR-P1 receptor CTLDs.

Methods

Protein preparation

DNA coding for the extracellular part of mNKR-P1A/C proteins (mNKR-P1A/C; to specify the origin of the receptor, the first letter of each organism name is used prior to the receptor name, i.e., m for mouse, r for rat, etc.), was amplified from the total mRNA isolated from the spleen of C57BL/6 mouse and subcloned into expression vector pET-30a (Novagen). Expression plasmids were transformed into *Escherichia coli* BL-21 (DE3) Gold (Stratagene). Bacteria were grown in LB medium, induction was performed with 0.1 mM isopropyl-β-D-thiogalactopyranoside and the induced culture was grown for 2 h. Proteins were refolded from inclusion bodies and were purified by HPLC ion exchange and gel filtration chromatography. The protein samples for experiments were concentrated at 10.4 mg/mL in a 15 mM Tris–HCl buffer with 150 mM NaCl, pH 8.0.

The rNKR-P1A/B proteins were prepared as described previously [4]. Proteins were dissolved in a 10 mM Tris-HCl buffer with 50 mM NaCl, pH 7.4, in concentrations of 9.1 mg/mL and 3.8 mg/mL, respectively. The concentrations were determined by Bradford assay [5].

Raman spectroscopy

Raman spectra of aqueous solutions of mNKR-P1A/C proteins were recorded in a standard 90° geometry on a multichannel instrument based on Spex 270 M single spectrograph with 1800 grooves/mm grating (Jobin-Yvon), a holographic notch-plus filter (Kaiser Optical Systems) and a liquid nitrogen cooled CCD detection system (Princeton Instruments) measuring 1340 pixels along the dispersion axis. The spectral resolution was approximately 5 cm^{-1}. Samples in a capillary micro-cell (5 μL inner volume) were excited with a 532.2 nm line (300 mW of radiant power per sample) NdYAG Verdi 2 laser (Coherent) and kept at 4 °C during all experiments using an external water bath (Neslab). The acquisition time for the spectra was 60 minutes. The spectra of temperature dependence each consisted of 5 min exposures. Temperatures in the 5–90 °C region were adjusted with 5 °C increments using the external water bath (Neslab) and equilibrated for 4 minutes before measurement. The wavenumber scale was calibrated with neon glow-lamp lines. Therefore, Raman frequencies of well-resolved bands are accurate to ±0.5 cm^{-1}.

The solutions of rNKR-P1A/B proteins were excited with a 514.5 nm Ar-ion laser Innova 300 (Coherent) using the same Raman spectrometer with the same experimental setup. Both spectra were accumulated for 600 min to produce traces of the highest quality.

Drop coating deposition Raman spectroscopy

The samples of mNKR-P1A/C (4 μL) were dialyzed (on Millipore filters 0.025 μm/white VSWP/13 mm) against deionized distilled water for 35 minutes. A 2 μL volume of protein solution, with an approximate concentration of 1 mg/mL, was deposited on a standard DCDR substrate SpectRIM™ (Tienta Sciences) consisting of a polished stainless steel plate coated with a thin layer of Teflon [6]. After air-drying at room temperature, approximately 20 minutes, Raman spectra were collected from “coffee rings” of former droplets [7] using a Raman microspectrometer HR800 (Horiba Jobin Yvon) with a 514.5 nm Ar-ion excitation laser (Melles Griot). A 50× microscope objective (N.A. 0.75, Olympus) was used to focus the 5 mW excitation laser to a diameter, approximately 1.5 μm, on the sample, and the spectra were integrated for 20 min using a 600 grooves/mm grating and liquid nitrogen cooled CCD detector (1024×256 pixels, Symphony). The spectrometer was calibrated using a band of Si-vibrations at 520.7 cm^{-1}. The spectral resolution was approximately 5 cm^{-1}.

The samples of rNKR-P1A/B proteins were treated in the same way as the mouse proteins. Nevertheless, the spectra of rNKR-P1A/B protein were integrated for 4 min using a 632.8 nm He-Ne laser excitation in the same setup as mentioned previously.

Infrared spectroscopy

Infrared spectra of mNKR-P1A/C proteins were recorded with a Bruker Vector 33 FTIR spectrometer using a

standard MIR source, a KBr beamsplitter and a DTGS detector. 5000 scans were collected with a Blackman-Harris 3-term apodization function at a spectral resolution of 2 cm^{-1}. Aqueous protein solutions were measured at room temperature in a CaF_2-cell with a 10 μm path length. Measurements in thermal dynamics were performed using a thermal cell holder BioJACK™ (BioTools). Temperature was adjusted from 5 to 90 °C with an increment of 5 °C using an external water bath (Neslab) and equilibrated for 4 minutes before measurements took place. 1000 scans were performed. Attenuated total reflection (ATR) FTIR measurements, 5000 scans, were realized by using an ATR-MIRacl™AG – single diamond horizontal ATR (Pike Technologies).

The rNKR-P1A/B proteins were measured using a Bruker IFS 66/S FTIR spectrometer equipped with an MCT detector. 4000 scans were collected with a Happ-Genzel apodization function at a spectral resolution of 4 cm^{-1}. The rest of the FTIR setup remained the same as the measurements for mNKR-P1 proteins. Spectral contribution of the buffer was corrected following the standard algorithm [8]. Spectrum of water vapors was subtracted and finally, all spectra were normalized.

Differential scanning calorimetry

The protein's stock solutions were diluted to the desired concentration – i.e., mNKR-P1A to 0.20 mg/mL and mNKR-P1C to 0.52 mg/mL. Calorimetric measurements were performed using the Model 6100 Nano II Differential Scanning Calorimeter – N-DSC II (Microcalorimetry Sciences Corporation). The samples were scanned from 5 to 90 °C at a heating rate of 1 °C/min under a constant excess pressure of 3 atmospheres. The appropriate DSCRun and CpCalc 2.2 software was used for data acquisition and analysis. After baseline subtraction of the buffer–buffer signal, the molar excess heat capacity function was obtained by dividing the protein concentration and cell volume (0.299 mL).

Sequence and phylogeny analysis

Sequences specified in Table 1 were aligned in ClustalX [21] and used for phylogeny analysis. For sequence analysis purposes, we used only their C-type lectin-like domains (CTLDs). Phylogenetic analysis was originally performed for protein sequences only; resultant trees however revealed too many polytomies. The number of polytomies was decreased by using nucleotide sequences for tree construction. Excluding Bayesian analyses, all protein and DNA sequences analyses were performed using Phylip 3.69 software [22] with the following algorithms: Neighbor-Joining [23], maximum parsimony [24], maximum likelihood [25] with assuming molecular clock and Fitch-Margoliash method assuming molecular clock (Kitsch method) [26]. Sequences were bootstrapped 1000× with the exception of the computationally most expensive method maximum likelihood, where sequences were bootstrapped 100×. To confirm that this value of bootstrapping was high enough, we constructed the trees with the same settings but bootstrapped 1100× or 110×. To calculate the distance matrix, and for the maximum likelihood method, we used the Jones-Taylor-Thornton matrix for protein sequences and the F84 matrix for DNA sequences. Where possible the sequences were jumbled 5×. In the neighbor-joining method the order of sequences was randomized. Probabilities of branch occurrence were calculated according to one of the most commonly used tests on the reliability of an inferred tree, Felsenstein's bootstrap test [27], which was evaluated using Efron's [28] bootstrap resampling technique. Resultant consensus cladograms were done using the 50% majority consensus rule. Bayesian analysis was performed using MrBayes 3.1.2 [29] with the same initial sequences used previously. Invariable gamma distribution and a GTR model was used to describe the parameters for the likelihood model, 150 000 for protein cycles and 100 000 for DNA Markov chain Monte Carlo cycles, with the state being swapped every 100×.

Homology modeling

Primary structures of mNKR-P1A, mNKR-P1C, rNKR-P1A, and rNKR-P1B extracellular domains were extracted from the database (Table 1). Templates were identified using BLAST [30] with matrix BLOSUM62. The identified homologs were almost the same for all seven models that were built. The template with highest identity to the target sequence was used. If some structures had the same identity, the one with the better resolution was used. Templates and identities used were: 1YPQ (C-type lectin-like domain of human oxidized low density lipoprotein receptor 1 (LOX-1) [31]) for mNKR-P1A/C (identity 31% / 32%), 1XPH (CD209 antigen-like protein 1 [32]) for rNKR-P1B (identity 33%), 3HUP (early activation antigen CD69 [33]) for mNKR-P1F (identity 33%), 1E87 (early activation antigen CD69 [34]) for mNKR-P1G (identity 35%) and 2BPD (beta-glucan receptor Dectin-1 [35]) for hNKR-P1 (identity 32%). Only in the case of rNKR-P1A was an alternative technique tried. The neighboring sequence from the neighbor-joining phylogenetic tree (1000× bootstrapped) was used, i.e., structure 3CAD (lectin-related NK cell receptor LY49G1 [36]). Alignment of template and modeled sequences used ClustalX. Alignments were checked to ensure that general features in sequence-conserved regions of CTLDs, were maintained as found in sequence multiple alignments of CTLDs. The resulting aligned sequences used for modeling are shown in

Table 1 Protein and nucleotide sequences used for phylogenetic and sequence analyses

Protein	Protein sequence	DNA sequence	Reference
rNKR-P1A^{WAG}	ABO15817	EF100677	Voigt et al. 2007 [9]
rNKR-P1A^{SD}	ABO15823	EF100683	Voigt et al. 2007 [9]
rNKR-P1A^{RNK}	ABO15821	EF100681	Voigt et al. 2007 [9]
rNKR-P1A^{F344}	AAA41710	M62891	Giorda et al. 1990 [10]
rNKR-P1A^{PVG}	ACJ47804	FJ416340	Kveberg et al. 2009 [11]
rNKR-P1B^{WAG}	ABO15818	EF100678	Voigt et al. 2007 [9]
rNKR-P1B^{BS}	AAQ08908	AF525533	Voigt et al. 2007 [9]
rNKR-P1B^{PVG}	ABA40404	NM_001040189	Voigt et al. 2007 [9]
rNKR-P1B^{SD}	ABO15824	EF100684	Voigt et al. 2007 [9]
rNKR-P1B^{RNK}	ABO15822	EF100682	Voigt et al. 2007 [9]
rNKR-P1B^{F344}	AAB01986	U56936	Dissen et al. 1996 [12]
rNKR-P1B^{TO}	AAQ11375	AF541943	Voigt et al. 2007 [9]
rNKR-P1F^{WAG}	ABO15819	EF100679	Voigt et al. 2007 [9]
rNKR-P1F^{SD}	ABO15825	EF100685	Voigt et al. 2007 [9]
rNKR-P1F^{F344}	CAA66111	X97477	Appasamy et al. 1996 [13]
rNKR-P1G^{WAG}	ABO15820	EF100680	Voigt et al. 2007 [9]
rNKR-P1G^{PVG}	ABA40405	DQ157011	Kveberg et al. 2009 [11]
rNKR-P1G^{DA}	ABA40406	DQ157012	Kveberg et al. 2009 [11]
mNKR-P1A^{B6}	AAA39822	M77676	Giorda et al. 1991 [14]
mNKR-P1A^{BALB}	ABB72025	DQ237927	Carlyle et al. 2006 [15]
mNKR-P1A^{SJL}	AAK39101	AF354261	Kung et al. 1999 [16]
mNKR-P1B^{B6}	AAK08512	AF338321	Carlyle et al. 2006 [15]
mNKR-P1B^{SJL}	AAK39099	AF354259	Kung et al. 1999 [16]
mNKR-P1B^{NOD}	BAE41390	AK169820	Carninci et al. 1999 [17]
mNKR-P1B^{BALB}	ABB72026	DQ23728	Carlyle et al. 2006 [15]
mNKR-P1C^{B6}	BAC37739	AK079743	Carninci et al. 1999 [17]
mNKR-P1C^{BALB}	ABB72027	DQ237929	Carlyle et al. 2006 [15]
mNKR-P1F^{B6}	NP_694734	NM_153094	Plougastel et al. 2001 [18]
mNKR-P1F^{BALB}	ABB72029	DQ237931	Carlyle et al. 2006 [15]
mNKR-P1F^{NOD}	BAE41516	AK170023	Carninci et al. 1999 [17]
mNKR-P1G^{B6}	ABC71751	DQ336141	Carlyle et al. 2006 [15]
hNKP-P1	NP_002249	NM_002258	Lanier et al. 1994 [19]
cNKR-P1	CAB55571	AJ245903	Kaufman et al. 1999 [20]

Fig. 8. Ten models were calculated for every protein and template using Modeller 9v4 [37]. The best model from every group was chosen by the distribution of amino acids in the Ramachandran plot and stereochemical *g*-factor (both calculated by Procheck [38]), Modeller distribution function and by visual inspection using SwissPDBViewer [39]. Secondary structure was determined by Procheck according to Kabsch and Sander [40].

Molecular dynamics

All selected models were minimized in SCP water solution with Gromacs 3.3.3 package [41], using virtual site hydrogens. The modified version of force field Gromos87 (usually called Gromacs) [42, 43] was used. Temperature was held at 300 K by separately connecting the protein and solution to the external temperature bath (t=0.1 ps) while the pressure was held at 1 bar by connecting to the pressure bath (t=0.1 ps). Algorithms SETTLE (for water) and LINCS (for protein) were used to restrict covalent bond length and long-range electrostatic interactions were calculated using the Particle-Mesh Ewald method. Optimization with steepest descent energy minimization was followed by solvent optimization using a time step of 1 fs for 10 ps. Counter ions were added to neutralize the simulation box and were consequently minimized for 10 ps using a time step of 1 fs. Finally, the protein was minimized for 20 ps using a time step of 2 fs. For production, the simulation runs were started using a time step of 5 fs. Root mean square deviation and radius of gyration analysis was

performed every 10 ns to check if the system reached equilibrium. The total production run time was 110 ns for mNKR-P1A, 10 ns for rNKR-P1A, 20 ns for rNKR-P1B and 100 ns for mNKR-P1C. The human and mouse models NKR-P1F/G were equilibrated for 10 ns.

After the dynamics run, all structures were minimized using the steepest descent algorithm. The representation of secondary structure types according to Kabsch and Sander [40] was calculated by Procheck software, and the root mean square fluctuation for each residue was calculated in Gromacs.

Results and discussion

Structural analysis of NKR-P1 proteins

Four NKR-P1 extracellular domains, rNKR-P1A and B and mNKR-P1A and C, were overexpressed in *E. coli*, refolded from inclusion bodies, and purified for spectroscopic analysis as described in Methods. Receptor domains were structurally analyzed by Raman, FTIR, and ATR-FTIR spectroscopy, including a novel drop-coating deposition Raman method that yields native-state spectra from very small solid samples [7]. Spectral assignments are given in Table 2, and secondary structure contents estimated using the pattern recognition least-squares method (LSA) are given in Table 3.

Figure 1 compares the FTIR spectra of mNKR-P1A and C domains. The similarity of the spectra is immediately obvious, indicating that mouse A and C proteins share the same fold with very similar secondary structure content. The slight differences in the spectra can be explained by their different amino acid compositions. For example, the second derivative band at 1517 cm^{-1} is connected with Tyr ring vibrations [54], and is more intense in mNKR-P1A that has five Tyr residues that in mNKR-P1C with only one Tyr. ATR-FTIR spectra of proteins were also measured for better resolution of subtle differences, and no further differences were detected (data not shown). The FTIR results are in excellent agreement with Raman spectroscopy data (Fig. 2 and Table 3). Solution Raman spectra (not shown) are also highly similar to drop-coated deposition spectra, which provide better signal-to-noise ratio.

The secondary structure estimates using the LSA method, which analyzes the amide I band in the case of Raman spectroscopy [49], and amide I and II bands in FTIR spectra [50], are not perfect for mNKR-P1A and C domains, suggesting a protein fold whose spectral pattern is not included in the LSA reference set [51]. Thus, the absolute percentages in the secondary structure content probably contain larger errors than estimated in Table 3. The fits for Raman spectral data are better. Nevertheless, both FTIR and Raman spectroscopy predictions differ by no more than 4% in each method for mouse A and C variants, and for each variant some predictions are in agreement within the margin of error. These results thus suggests only a small difference in the content of α-helices and β strands of about 10% for the two mouse domains.

Negative bands are observed in the Raman difference spectrum between mouse variants A and C in the region reporting on S-S-bridge conformations (Fig. 2). The presence of these intense and narrow vibration bands

Table 2 The assignment of Raman vibrational bands distinguishable in the spectra of rat NKR-P1A protein (see Fig. 4)

Frequency (cm^{-1})	Assignment	Notes
509	νS–S in GGG	[44]
539	νS–S in TGT	[44]
572	Trp	[44]
623	Phe	[45, 46]
642	Tyr	[45, 46]
697	νC–S P_N, P_H	[44]
760	Trp W18	[45–47]
828	Tyr doublet, Phe	[44–47]
853	Tyr doublet, Ile	[44–47]
878	Trp W17, Ile, Val	[45–47]
894	Lys, Ala	[45, 46]
937	Lys, Val, Leu	[45, 46]
959	Lys, Leu	[45, 46]
988	Ile	[45, 46]
1005	Phe	[44–46]
1011	Trp W16 shoulder	[44–47]
1033	Phe, Tyr	[44–46]
1062	Lys, Ala, Phe	[45, 46]
1077	Lys, Phe	[45, 46]
1127	Ile, Val, Leu, Trp	[45, 46]
1158	Ile, Val	[45, 46]
1175	Tyr, Phe	[45, 46]
1206	Phe, Tyr, Trp shoulder	[45, 46]
1237	amide III	[45, 46]
1315	amide III, Lys, Val, Ile	[45, 46]
1339	Trp doublet	[44–47]
1358	Trp doublet, Val, Phe	[44–47]
1404	νCO_2^- of Asp, Glu	[47]
1425	Trp shoulder	[45, 46]
1449	δCH_2, δCH_3, Lys, Ile, Leu	[45, 46]
1462	δCH_2, Ala, Ile, Val, Leu, Trp, Tyr	[45, 46]
1551	Trp W3	[48]
1579	Trp, Phe	[45, 46]
1607	Phe, Tyr Y8b	[45–47]
1620	Tyr Y8a, Trp	[45–47]
1671	amide I	[44–47]

Table 3 An estimation of secondary structure content of mNKR-P1A/C and rNKR-P1A/B proteins in aqueous solution by pattern recognition least-squares method (LSA) analyzing the amide I band in the case of Raman spectroscopy [49], and amide I and II region in the case of FTIR [50] (using the spectral set from ref. [51]). Secondary structure estimation is compared with modeled structures. Given standard deviations are calculated as standard deviations of the used reference set, therefore they do not reflect the quality of the fits

	Mouse NKR-P1A			Mouse NKR-P1C		
Structure	Raman	FTIR	Model	Raman	FTIR	Model
α-helix	28±5	22±10	15	26±5	20±10	20
β-sheet	38±4	26±9	39	37±4	24±9	37
β-turn	17±2	14±4	10	18±2	14±4	14
β-bend	—	14±4	18	—	18±4	10
Unordered	17±2	24±6	18	19±2	24±6	19
	Rat NKR-P1A			Rat NKR-P1B		
Structure	Raman	FTIR	Model	Raman	FTIR	Model
α-helix	15±5	13±10	18	19±5	25±10	25
β-sheet	50±4	32±9	40	48±4	26±9	33
β-turn	19±2	13±4	15	19±2	10±4	12
β-bend	—	16±4	11	—	14±4	12
Unordered	15±2	26±6	16	14±2	25±6	18

reflects a more well-defined conformation of S–S bridges in mNKR-P1C and thus lower flexibility. Thus, the C variant is probably more rigid than the A variant, in agreement with the higher stability detected in calorimetric data as discussed below.

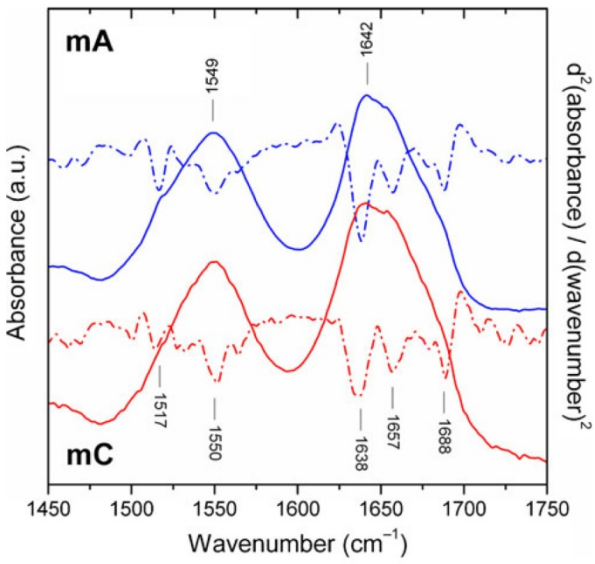


Fig. 1 A comparison of FTIR spectra from mouse NKR-P1A protein (mA) with its mouse NKR-P1C variant (mC) in the region of amide I and amide II bands. The dash-dot lines represents second derivative of the spectrum (smoothed by Savitski-Golay function at 15 points, i.e. ca. 14 cm^{-1}). The peak positions are labeled for NKR-P1A protein. Slight shifts are therefore visible in the second derivative of FTIR NKR-P1C spectrum. Two major bands, observed at 1642 and 1549 cm^{-1}, are assigned to amide I and amide II vibrations of the backbone chain, respectively. The second derivative, which can identify overlapping components, reveals a high content of extended β-sheets by a strong negative band at 1638 cm^{-1} [52]. The second component of β-sheets at 1688 cm^{-1} suggests that β-sheets are antiparallel [53]. The negative band at 1657 cm^{-1} corresponds to α-helices and an unordered structure [52] which points towards an α/β protein fold. The second derivative band at 1688 cm^{-1}, with its shoulder at 1678 cm^{-1}, is connected with high content of loops and turns

More distinctive differences can be seen between the FTIR spectra of rNKR-P1A and B domains (Fig. 3). The band positions of the rat A variant differ only slightly more than the differences between the two mouse variants, and therefore the rat A variant can be structurally clustered with both mouse variants; rNKR-P1B is the most distinctive protein within the measured group. Raman difference spectroscopy (Fig. 4) adds detail to these distinctions, with significant differences between A and B variants in the amide I region, ca. 1671 cm^{-1} and amide III, ca. 1237 cm^{-1}. In the amide III region, the positive band, ca. 1230 cm^{-1}, corresponds to an excess of β-structures in rNKR-P1A relative to the B variant. The negative band at 1280 cm^{-1} corresponds to lower α-helix content in the A variant when compared with the B variant. The aromatic amino acid composition of rNKR-P1A/B domains are the same; thus the Raman difference spectrum in Fig. 4 reflects mostly secondary structure differences between the proteins. Comparing all Raman spectra, rNKR-P1B protein is again the most distinctive, with mouse A and C variants very similar, and rat variant A being somewhere in between.

Thermal dynamics

Differential scanning calorimetry showed an irreversible denaturation transition for mNKR-P1A and C domains corresponding to a single, noncooperative thermal transition, with melting temperatures of 69 °C for mNKR-P1A and 72 °C for the C variant (data not shown). Thermodynamic parameters for the transitions, determined as described in Methods, were the same for both proteins, ΔH=48 kcal·mol^{-1}, ΔS=0.14 kcal·K^{-1}·mol^{-1}. Therefore the mNKR-P1C domain is slightly more thermostable than the A variant.

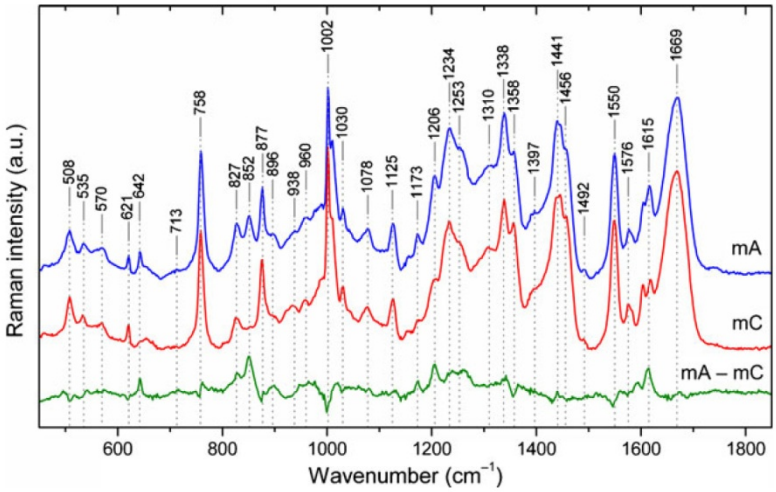


Fig. 2 A comparison of mouse NKR-P1A DCDR spectra (mA) with its mouse NKR-P1C variant (mC); (mA – mC) denotes appropriate spectral difference. In the amide I region at 1669 cm^{-1}, corresponding mostly to stretching νC_{α}=O skeletal vibrations (for band assignments see Table 2), small negligible differences between mNKR-P1A and C protein variants can be found (mA – mC difference). This is supported by the skeletal NH-vibration region of amide III at 1310 cm^{-1}, where the difference is almost zero between mouse proteins. The band at 508 cm^{-1} corresponds to the lowest free energy *gauche–gauche–gauche* (GGG) conformation of S–S bridges, the most obvious conformation in the protein structures. The band at 535 cm^{-1} reflects unusual conformation that is close to *trans–gauche–trans* – TGT (pure conformation is characterized by vibrations at 540 cm^{-1}, whereas *gauche–gauche–trans* (GGT) has a band at ca. 525 cm^{-1}), which has the highest free energy [44]. The intensity ratio I_{508} / I_{535} clearly indicates that two S–S bridges have GGG conformation and one nearly has TGT conformation. Direct correlation between the Raman frequency of W3 mode and the absolute value of the torsional angle $\chi^{2,1}$ is known [48]. Its position at 1550 cm^{-1} corresponds to a 93° angle | $\chi^{2,1}$| as an average angle of six Trp in mNKR-P1A/C proteins. The negative difference band at 1546 cm^{-1} alludes to prevalence on lower angles in the C variant. The Trp W17 mode at 877 cm^{-1} indicates moderate NH–hydrogen bond donation in both A and C variants [55]. The Trp Fermi doublet intensity ratio I_{1358} / I_{1338} ca. 0.88 is a sensitive marker of the amphipathic environment of the aromatic ring [55] and in our case indicates a hydrophilic environment of all Trp in both variants. The two bands at 827 and 852 cm^{-1} are assigned to the Tyr (Y1 + Y1a) Fermi resonance doublet. Its intensity ratio I_{852} / I_{827} is used as an Tyr environment indicator [56]. The value 1.0 of mNKR-P1A protein corresponds to moderate Tyr H-bonding, mostly at the surface of the protein. Mouse NKR-P1C protein has only one Tyr with an intensity ratio of 0.75, which may suggest that phenolic OH of this Tyr is a stronger hydrogen bond donor than those placed on the surface of the protein

Two-dimensional correlation spectroscopy (2DCoS) was applied to FTIR and Raman spectra to add structural detail to the thermal unfolding transitions of mNKR-P1A/C proteins. The aim of this technique is to identify in-phase and out-of-phase correlations between spectral intensity variations occurring at different wavenumbers that are induced by external perturbation of the studied system; a generalized formalism for 2DCoS can be found in ref. [57–59]. The main advantage of 2DCoS is that it allows enhancement of spectral resolution by spreading overlapping bands over a second dimension. In addition, sign analysis of the correlation peaks in the 2D maps may improve band assignment and permit establishment of a sequence of events during the perturbation process [57–59].

Synchronous 2D correlation FTIR spectra (Fig. 5 left) suggest slightly lower stability in the A variant relative to C, in agreement with calorimetric data, and suggest this transition is correlated with exposure of hydrophobic regions in β-strands. The Raman synchronous spectrum is significantly less complicated (Fig. 6 left) and reports on the same denaturation process, consistent with the FTIR results. The asynchronous 2DCoS spectra (Figs. 5 and 6 right) reveal sequential, but not coincidental, spectral changes. The asynchronous spectrum has no autopeaks and consists exclusively of crosspeaks that are antisymmetric in regard to the diagonal line. Asynchronous cross peaks develop only if the intensities of two spectral features change and are out of phase with each other. If the asynchronous peak sign becomes positive then the intensity change at the given wavenumber on the *x*-scale occurs predominantly before it is connected with the 2DCoS *y*-scale and vice versa. This sign rule is reversed if the synchronous correlation intensity at the same coordinate becomes negative.

Application of these so called Noda rules [57–59] to mNKR-P1C second derivative FTIR spectra indicate the following sequential order of structural changes. Changes in β-turns occur first, followed by an increase in β-sheet content, then changes in α-helices followed by changes in loops. These structural changes precede emergence of β-aggregated structures, which is followed by changes in β-turns. The sequential order in the mNKR-P1A variant is the

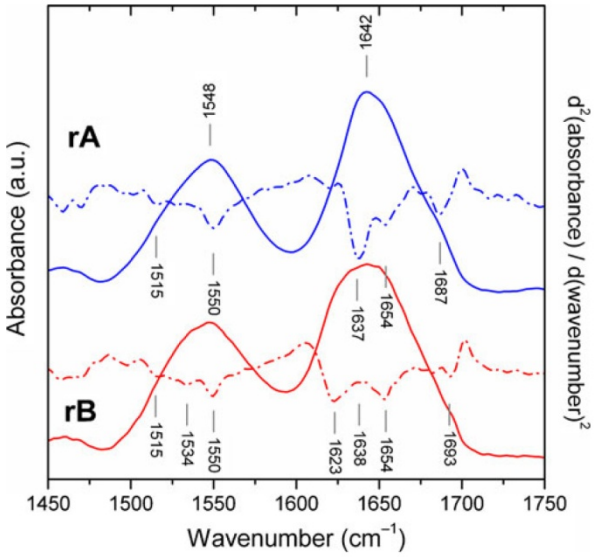


Fig. 3 A comparison of rat NKR-P1A FTIR spectra (rA) with its rat NKR-P1B variant (rB) in the region of amide I and amide II bands. The dash-dot lines represents second derivative of the spectrum (smoothed by Savitski-Golay function at 15 points, i.e. ca. 21 cm^{-1}). Although the rNKR-P1B spectrum was partially affected by the low stability of the sample which caused partial aggregation (resolved by second derivative bands at 1623 and 1693 cm^{-1}) [52], higher amounts of helical structures (second derivative band at 1654 cm^{-1}) in comparison with antiparallel β-sheets (negative bands at 1638 and shoulder at 1682 cm^{-1}) can be seen same, although the emergence of β-aggregates was not observable. The sequential order of secondary structure changes from FTIR is in agreement with that from Raman data. Asynchronous analysis of Raman data shows that changes in Tyr/Phe vibration at 1600 cm^{-1} precede changes in β-turns, and the increase of β-sheet content is followed by changes in Trp/Phe vibrations at 1590 cm^{-1}.

Heterospectral 2DcoS was used because it can reveal correlations between different spectral regions or even between two different spectroscopic techniques and may help to detect vibrations of a similar nature or connections with the same process [59]. Only synchronous heterospectral 2DCoS Raman spectroscopy was used to permit direct correlation of secondary structure elements with specific residue types, and because Raman spectra contain more residue-specific spectral information than FTIR spectra. The synchronous spectra display only cross peaks (Fig. 7) because autopeaks are generated only by the correlation of the same bands as in Figs. 5 and 6. Correlations were investigated in the changes of the secondary structure, represented by the Raman amide I region (1580–1700 cm^{-1}), and the regions where aromatic side chains have intense Raman bands. The strong correlations reported in Fig. 7 suggest that one or more Trp residues experience major changes in their surroundings during the early stages of thermal denaturation.

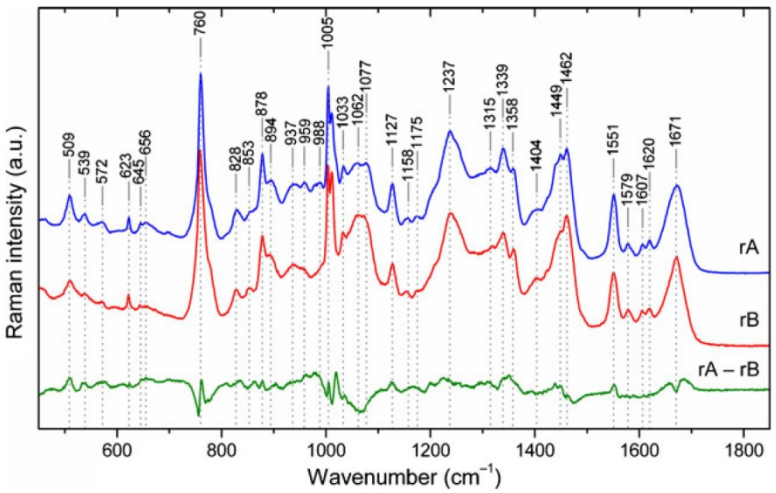


Fig. 4 A comparison of rat NKR-P1A DCDR spectra (rA) with its rat NKR-P1B variant (rB); (rA – rB) denotes appropriate spectral difference. It is clearly visible in the disulfide bridges stretching vibrations region (the bands at 509 and 539 cm^{-1}), that two S–S bridges are in GGG conformation whereas one is in pure TGT conformation. In comparison with mouse protein A/C variants, rat variants show a higher tension on this TGT bridge. The S–S bridges in rNKR-P1B are much more flexible than in the A variant (exhibiting broader and less intense bands) which points to a lower stability of this protein. Only small differences exist between all mouse and rat proteins when comparing the absolute value of the torsional angle $\chi^{2,1}$ that is characterized by the band at 1551 cm^{-1}, which corresponds to an average 95° angle of six Tyr. Nevertheless, there is a positive peak, also visible within the Raman difference spectra, that is reflecting higher flexibility in the rNKR-P1B protein. In the case of mouse A/C variants, other prominent Tyr bands (878, 1339, 1358 cm^{-1}) reflect the same phenomena as mentioned above whereas each of these vibrations shows higher flexibility in the rNKR-P1B protein. The rat proteins have only one Tyr whose ratio 1.0 of 828 and 853 cm^{-1} bands corresponds to moderate H-bonding which is characteristic for the position on the protein surface

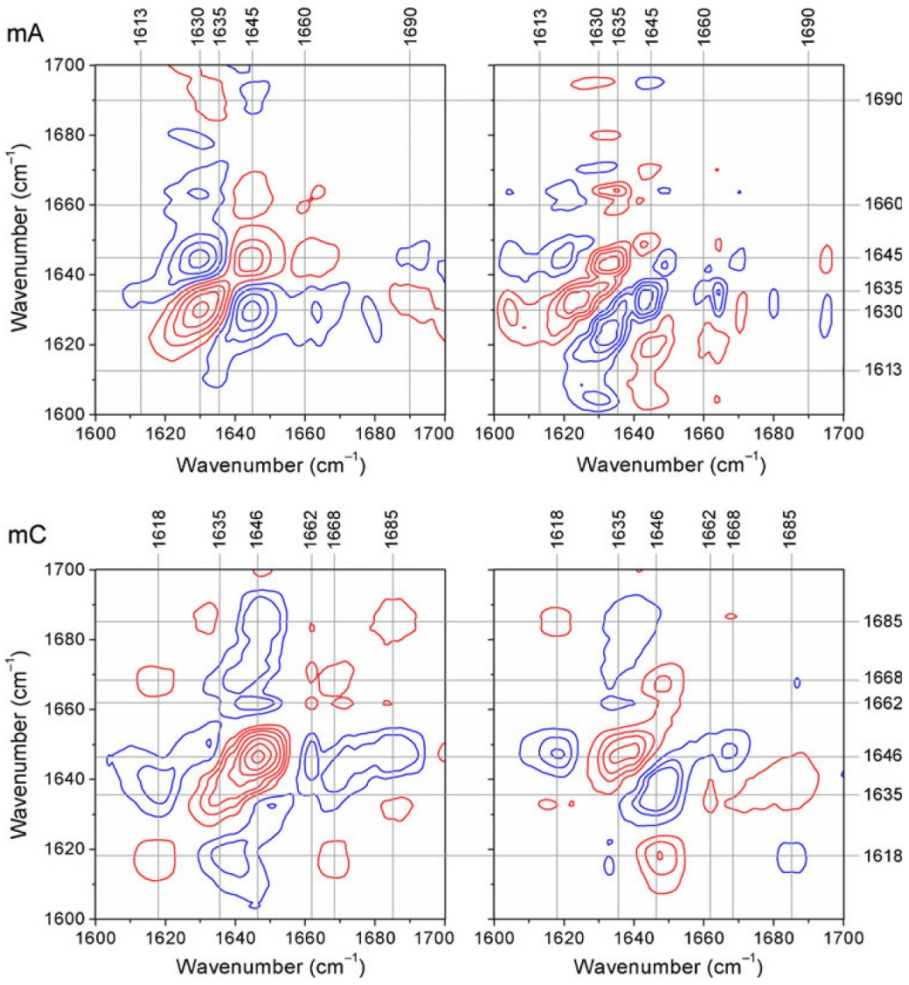


Fig. 5 Synchronous (left) and asynchronous (right) 2D correlation FTIR spectra of thermal dynamics from 5 °C to 90 °C (with 5 °C increments) of mouse NKR-P1A (mA) and NKR-P1C (mC) proteins. The white and red peaks are positive while the grey and blue peaks are negative. The synchronous second derivative 2DCoS spectra from FTIR spectra of mNKR-P1A/C, under thermal perturbation from 5 °C to 90 °C, exhibit several positive diagonal peaks, so called autopeaks (see left side). The diagonal spectrum of the synchronous 2DCoS mathematically corresponds to autocorrelation function. Thus, any region of the spectrum which changes intensity to a great extent, under the thermal perturbation, will show strong autopeaks, while those remaining almost constant develop no autopeaks. In the case of protein mNKR-P1C, thermal denaturation is connected with changes of aggregate β-sheets (1618 cm^{-1}), native β-sheets (1635 cm^{-1}), α-helices and unordered structures (1646 cm^{-1}), loops (1668 cm^{-1}) and β-turns (1685 cm^{-1}) [52] as demonstrated by its autopeaks. Synchronous 2DCoS describes in-phase intensity variations, i.e., structural changes, so it is questionable if the correlation is positive or negative. This question can be answered by off-diagonal peaks, so called cross peaks. While the autopeaks sign is always positive, the cross peaks sign can be either positive or negative. The synchronous cross peak sign becomes positive if the spectral intensities at the two coordinates are simultaneously increasing or decreasing. On the other hand, a negative sign at the intensity of the cross peak indicates that one of the spectral intensities is increasing while the other is decreasing. Thus, the cross peaks correspond to a simultaneous decrease of α-helices and unordered structures with an increase of β-sheets, loops and β-turns whereas the decrease of β-sheets is then connected with the increase of β-aggregated structures. Similar structural changes can be found in synchronous FTIR 2DCoS in protein mNKR-P1A, however its transition to β-aggregates was not observable because the process was too fast

The structural changes occurring in NKR-P1 proteins during the increase of temperature can be interpreted tentatively as follows. Some large flexible part with high content of β-turns is rearranged first. Trp residues appear to be involved in significant dynamical behavior during these early stages of denaturation. Next, the β-sheet content increases before or during continuous decrease of α-helix content. These changes lead to the exposure of a hydrophobic region, perhaps containing Tyr and/or Phe residues, leading to β-aggregates followed by changes in β-turns.

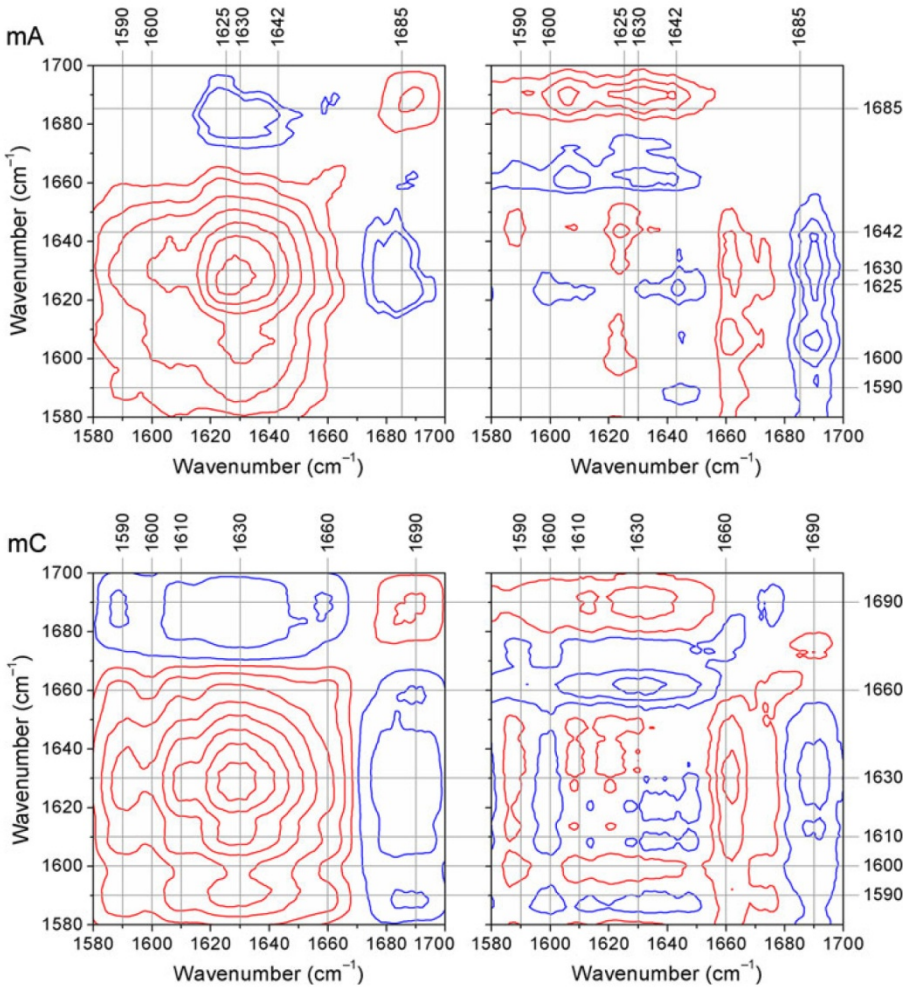


Fig. 6 Synchronous (left) and asynchronous (right) 2D correlation Raman spectra of thermal dynamics from 5 °C to 90 °C (with 5 °C increments) of mouse NKR-P1A (mA) and NKR-P1C (mC) proteins. The white and red peaks are positive while the grey and blue peaks are negative. The decrease of α-helix content (the band at 1635 cm^{-1}) is connected with an increase of β-sheet and β-turn structures (1685/1690 cm^{-1} for mNKR-P1C and mNKR-P1A, respectively). This structural change is connected with environmental changes in Trp, Tyr and Phe residues (corresponding bands at 1590, 1600 and 1610/1625 cm^{-1})

Modeling

Homology modeling and molecular dynamics

Candidate template sequences were identified using BLAST [30] and pair-aligned with modeling target sequences using ClustalX as described in Methods. The template structure with highest sequence identity to each target was used for modeling; identities ranged from 31 to 35%, and the resulting aligned sequences used for modeling are shown in Fig. 8. Template structures used were: 1YPQ (C-type lectin-like domain of human oxidized low density lipoprotein receptor 1 (LOX-1) [31]) for mNKR-P1A/C, 1XPH (CD209 antigen-like protein 1 [32]) for rNKR-P1A/B, 3HUP (early activation antigen CD69 [33]) for mNKR-P1F, 1E87 (early activation antigen CD69 [34]) for mNKR-P1G, and 2BPD (beta-glucan receptor Dectin-1 [35]) for hNKR-P1. These template structures are similar to each other as judged by C_α root-mean-square deviations ranging from 0.94 to 1.50 among them, and all have functions close to those of the NKR-P1 proteins [61]. C-type lectin-like NK receptors are dimeric in their nature, but as the mode of dimerization is not conserved within this receptor family, we did not attempt to model NKR-P1s as dimers. Moreover, in native receptors covalent linkage by one or more disulfide bridges in the so-called "stalk region" close to the cell membrane is involved in dimerization. This region is usually omitted from soluble recombinant receptor domain constructs and thus is also missing in available crystal structures. Finally, all recombinant NKR-P1s prepared in this study behaved like a monomeric species during their purification (data not shown), pointing out for rather weak dimer formation within the NKR-P1 family, however, we cannot exclude that cooperativity between monomers will occur upon ligand binding. For consistency, the numbering of sequences begins at residue 89 in all sequences, and specific residue numbers discussed in the text are those of

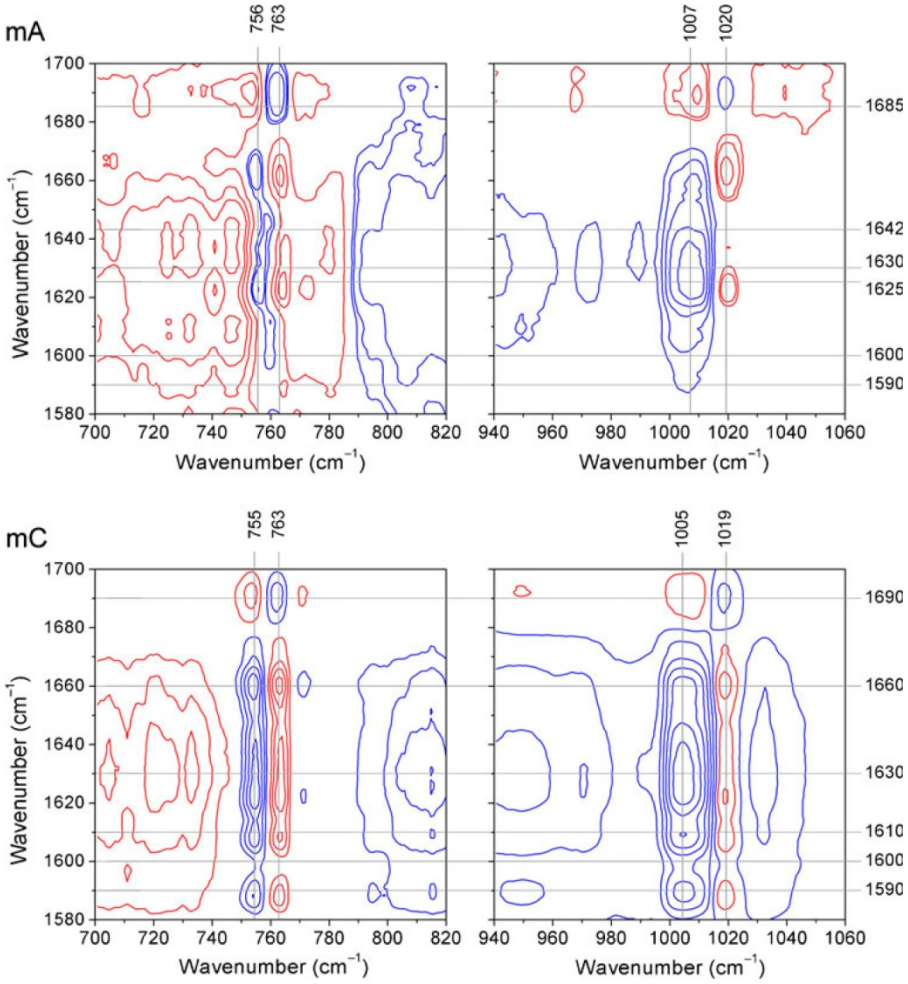


Fig. 7 Heterospectral synchronous 2D correlation analysis Raman spectra of thermal dynamics from 5 °C to 90 °C (with 5 °C increments) of mouse NKR-P1A (mA) and NKR-P1C (mC) proteins. The dynamics of two different regions against the amide I region (1580–1700 cm^{-1}) was followed. The white and red peaks are positive while the grey and blue peaks are negative. An intense search through the entire Raman spectrum did reveal very strong correlations between the amide I region and Trp bands at 760 and 1011 cm^{-1} for both mNKR-P1 variants. It is well known that intensity changes of these Trp Raman bands are connected with amphipathic environment of the indol ring [60]. Thus, 2DCoS reflects changes in hydrophobicity in the Tyr environment by the band shifts from 756 cm^{-1} to 763 cm^{-1} and from 1007 cm^{-1} to 1020 cm^{-1} for mNKR-P1A and similarly for mNKR-P1C. Moreover, the sign for 2DCoS bands is reversed in the amide I region, ca. 1690 cm^{-1}, which corresponds to β-turn structures in proteins

the mNKR-P1A sequence. Ten models were calculated for every sequence-template pair using Modeller 9v4 [37]. For further analysis the best model from every group was chosen as described in Methods, and refined by at least 10 ns of molecular dynamic simulations at room temperature in explicit solvent.

All models share the same basic α/β fold represented in Fig. 9 by rNKR-P1B. Helices surround a beta core composed of long antiparallel β-strands 2 and 8 that form a central 'pillar' flanked at one end by short strands 1 and 3 and at the other end by a small antiparallel sheet formed by strands 4–7. Although β-strand 3 is very short, its residues are highly conserved and they adopt very similar positions in all models (Table 4). The lengths of secondary structure elements (Table 5) reveal that a conserved number of residues forms strands 2 and 8 of the central pillar except in mNKR-P1F that is 1–2 residues shorter. These strands define the height of the receptor on the outside surface of the cell membrane. The models contain the three disulfide bonds identified in experimental data [62], except for mNKR-P1C where Cys122 is replaced by Ser.

Although the core of the protein is nearly the same for all seven models (C_α rmsd <2 Å for the beta core defined above), the number and orientation of helices as well as the topological organization of secondary elements vary somewhat among the models, defining four groups (Fig. 10). Mouse NKR-P1A/C models are more similar to each other than rat models are to each other. In addition to the topology differences, the prominent loops anchored by small antiparallel sheets differ in their arrangement among models, making rmsd comparisons meaningless for these segments. In all cases one small sheet (residues Ala127, Tyr128 and Leu129 in mNKR-P1A) anchors one loop of ~5 to 7 residues (residues Met130 to Gln135 in mNKR-P1A, green loop in Fig. 10, first half of region III in Fig. 12) and another small sheet (residues Trp165, Lys166 and Trp167 in mNKR-P1A) anchors one long loop of ~19 residues (e.g., residues Arg168 to Asp187 in mNKR-P1A, blue loop

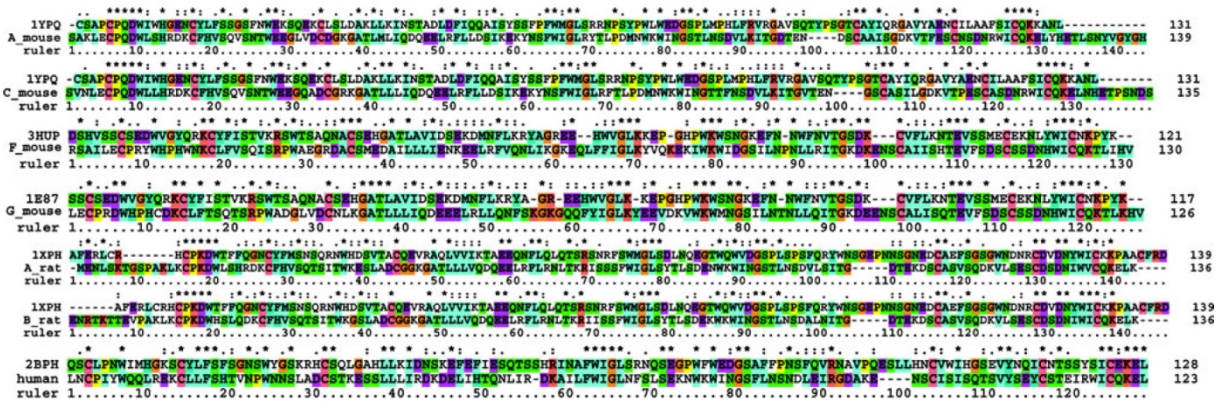


Fig. 8 Alignments for homology models in the following order; mouse NKR-P1A, rat NKR-P1A, rat NKR-P1B, mouse NKR-P1C, mouse NKR-P1F, mouse NKR-P1G and human NKR-P1

between sheets 5 and 6 in Fig. 10, region V until beginning of region IVb in Fig. 12). In rNKR-P1B and hNKR-P1 the long loop contains an additional short helix that packs on one side of the core together with helix B. In rNKR-P1B the connectivity of the loops to the small sheet also differs.

These differences in the spatial arrangement of the loops among the models are reflected also in the behavior of each model in molecular dynamics simulations. In all simulations the protein core is stable, with C_α rmsds reaching plateau values of ~1.3 to 2.5 Å (data not shown). The root mean square fluctuations for C_α calculated from the last nanosecond of the equilibrated part of each trajectory clearly identify the long loop region as very flexible (Fig. 11). Both ends of the loop, immediately after the N- and C-terminal anchoring residues, are extremely flexible. The root mean square fluctuations of the other residues in the loop are slightly higher than the average for the whole protein, although the structure within the loop is stable in the time period that could be simulated. The N-terminal anchoring residues are the highly conserved WKW sequence motif (residues Trp165, Lys166 and Trp167 in mNKR-P1A). The loop occasionally folds back onto the protein surface, where it forms a hydrophobic interaction with the two tryptophan residues. Together these results suggest that the loop region is stably anchored to the core, but can adopt alternative positions relative to the core.

Fig. 9 Detail view on the rat NKR-P1B model

Sequence analysis

The 33 available NKR-P1 CTLD domain sequences (Table 1) were aligned in ClustalX [21]. Seven conserved regions are identifiable (Fig. 12). Among these regions, conservation among orthologs is highest in the beta core and lowest in the two loops, the smaller loop corresponding to region III and the large loop corresponding to region V. In region III the chemical properties of loop residues are preserved despite the sequence variations. In contrast, substitutions in extended region V (L160 to T183) cause significant changes in the chemical character of some loop residues among natural NKR-P1 variants, leading to two groups typified by NKR-P1 subfamilies A-D and F/G. The NKR-P1A/D group presents L160, N164, T171, K178 and T183, whereas the NKR-P1F/G group substitutes a polar residue (Glu or Gln) for Leu160, a beta-branched residue (Thr, Val, or Ile) for Asn164, a more hydrophobic residue (Ile or Val) for Thr171, a Ser or Lys residue for Arg178, and a more polar residue (Asn, Asp, or Glu) for Thr183. The results of *in vivo* binding experiments with C-type lectin receptor (Clr) isoforms [11] suggest that these sequence differences may be related to ligand specificity: rNKRP1A and rNKRP1B bind to Clr11 only; rNKRP1F and rNKRP1G bind to Clr2, Clr6, and Clr7 but not to Clr11, and rNKRP1F also binds to Clr3 and Clr4. Thus, the long adaptable loop of the C-type lectin-like domain may encode its ligand specificity.

Phylogenetic analysis

Phylogenetic analysis was originally performed with only the protein sequences of the C-type lectin-like domains,

Table 4 Amino acids (their primary structure number) that form α-helices or β-strands in every model that was calculated in Procheck. In mouse NKR-P1F, β-strand 4 is divided into two regions. For this reason there are two numbers in the relevant column

	I	II	A	III	B	IV	C	V	VI	VII	VIII
Mouse A	100 – 102	104 – 114	119 – 127	129 – 131	135 – 146	152 – 159		167 – 169	188 – 194	197 – 203	207 – 215
Rat A	97 – 101	103 – 113	114 – 125	128 – 130	134 – 145	150 – 159		164 – 168	188 – 194	196 – 200	206 – 215
Rat B	100 – 102	104 – 114	115 – 126	129– 131	135 – 149	152 – 155	175 – 181	182, 183	190 – 194	197 – 201	207 – 215
Mouse C	100 – 102	104 – 114	115 – 126	129 – 131	135 – 149	152 – 159		167 – 169	188 – 194	197 – 201	207 – 215
Mouse F	98 – 102	105 – 114	115 – 126	129 – 131	135 – 146	152 – 154/ 156 – 158		167 – 169	190 – 192	199 – 201	207 – 213
Mouse G	98 – 102	104 – 114	115 – 126	129 – 131	135 – 146	152 – 159		167 – 169	188 – 194	197 – 203	207 – 215
Human	98 – 102	105 – 115	116 – 125	129 – 131	135 – 147	153 – 160	175 – 180	165 – 169	189 – 194	197 – 201	205 – 213

starting at position 89 for mouse NKR-P1A. Although five methods gave similar results, cladograms contained a large number of polytomies. DNA nucleotide sequences were then applied to bring in additional information for tree construction. Cladograms based on DNA sequences were constructed with several different approaches as described in Methods, leading to nearly identical results. In every case, the major branches of the cladogram were the same, as confirmed by bootstrapping. Bootstrapping failed only when the arrangement of species within the clade was different. This robustness of the final result supports the principal correctness of the tree. The resultant cladogram obtained by the maximum-likelihood method is shown in Fig. 13. Where some methods suggest polytomy or further branching, both possibilities are drawn.

The cladogram displays three major branches. The first is composed of human and chicken NKR-P1s. The second main branch is composed of rodent NKR-P1s from the telomeric part of the NK gene complex. In this branch the sequences are first branched according to the receptor subfamily and then according to the species. The third major group consists of rodent receptors from the centromeric part of the NK gene complex. Here, the sequences are ordered first according to species and then according to receptor subfamily. The most undefined part of the cladogram includes the rat NKR-P1As, as some methods even create pentatomy.

Table 5 The number of residues in secondary structure features. In mouse NKR-P1F, β-strand 4 is divided into two regions. For this reason there are two numbers in the relevant column

	1	2	A	3	B	4	C	5	6	7	8
Mouse A	2	11	9	3	12	8		3	7	7	9
Rat A	5	11	12	3	12	10		5	6	5	9
Rat B	3	11	11	3	12	4	7	2	5	4	9
Mouse C	3	11	12	3	15	8		3	6	5	9
Mouse F	5	10	12	3	12	3+3		3	3	3	7
Mouse G	5	11	12	3	12	8		3	7	7	9
Human	5	11	10	3	13	8	6	5	6	5	9

Comparison of modeling and spectroscopic results

The extent of predicted α-helical structure in rNKR-P1B agrees closely with experiments, especially with the data from FTIR. In the case of mNKR-P1A/C, more α-helical structure is detected in the spectra than is present in the models. One likely explanation that would be in quantitative agreement with the data is that short α-helix C, which is present in human and rat NKR-P1B models, is also present in both mouse structures. In the case of β-strands, the algorithm used to determine hydrogen bonds in models uses an average cutoff value, whereas in nature this cutoff is not so sharp, and it also depends on the β-strand environment and other factors. The agreement between the spectroscopic data and the models is particularly strong for beta turns, especially when measured by FTIR, whereas Raman spectra give slightly higher turns content. The phylogeny depicted in Fig. 13 is also consistent with the vibrational spectra, which group mouse A/C variants together as most distinct from rNKR-P1B, and with rat variant A being somewhere in between.

Disulfide bond conformations for Cys122-210 and Cys94-105 are unambiguously in a GGG conformation in all computational models. Therefore these two bridges can be assigned to the two clear GGG conformations determined from the band at 509 cm^{-1} (Fig. 2) representing the disulfide bridges region of stretching vibrations. The third disulfide bridge between Cys189-202 in the models adopts conformations in between the more rare GGT and TGT conformations, and could be assigned to the TGT conformation observed experimentally (Fig. 2).

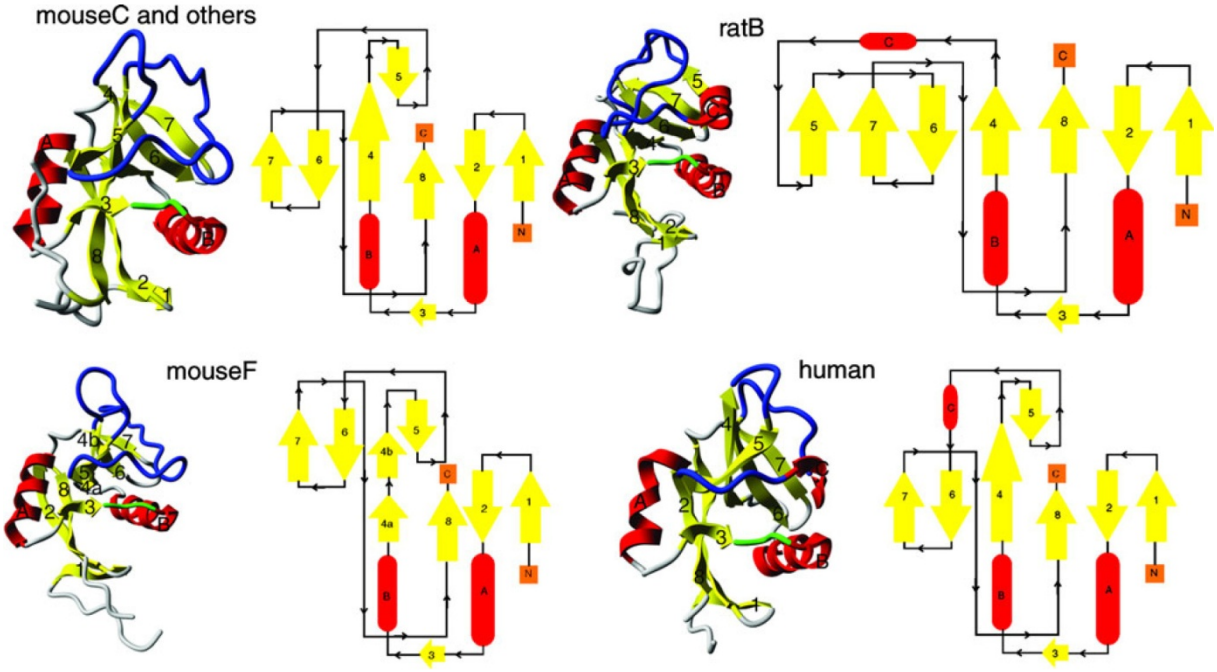


Fig. 10 Structures and topologies of resultant models. The regions discussed in the text are depicted here. The same fold like mouse NKR-P1C receptor is also adopted by mouse NKR-P1A/G and rat NKR-P1A. Mouse NKR-P1F, rat NKR-P1B and human NKR-P1 folds are unique

The inferred conformational adaptability of the long loop region finds support in the spectroscopic data as well. Opening of the loop as suggested by its behavior in molecular dynamics simulations would expose the two anchoring tryptophan residues to solvent, in agreement with the early shifts observed for tryptophan residues in the experimental thermal dynamics data.

Conclusions

MD-refined homology models of rat NKR-P1A/B, mouse NKR-P1A/C/F/G and human NKR-P1 receptor C-type lectin-like extracellular domains were generated in order to understand evolution of the structural features determining their functions. The rat NKR-P1A/B and mouse NKR-P1A/C domains were structurally analyzed by Raman and FTIR spectroscopy. Combined protein and DNA sequence analysis divided NKR-P1s into three major phylogenetic groups.

The derived models agree well with the Raman and infrared spectra not only in terms of overall secondary structure content, but additionally the models capture key structural and dynamic features consistent with thermal dynamics and MD simulations. Two sequence regions that are conserved in all analyzed NKR-P1s, and thus under strong evolutionary pressure, define the major phylogenetic branches, including the distinction between the NKR-P1A/D and the NKR-P1F/G subfamilies. One of these regions is the long loop present in nearly every CTLD; it is alternatively arranged in the different models, and is predicted to contain only a minor portion of secondary

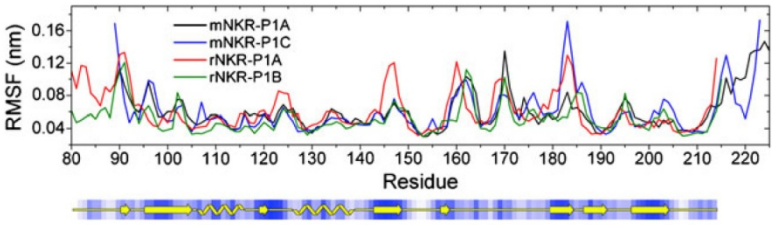


Fig. 11 Root mean square fluctuation of amino acid residues during the production phase of the simulation is represented by C_α carbon

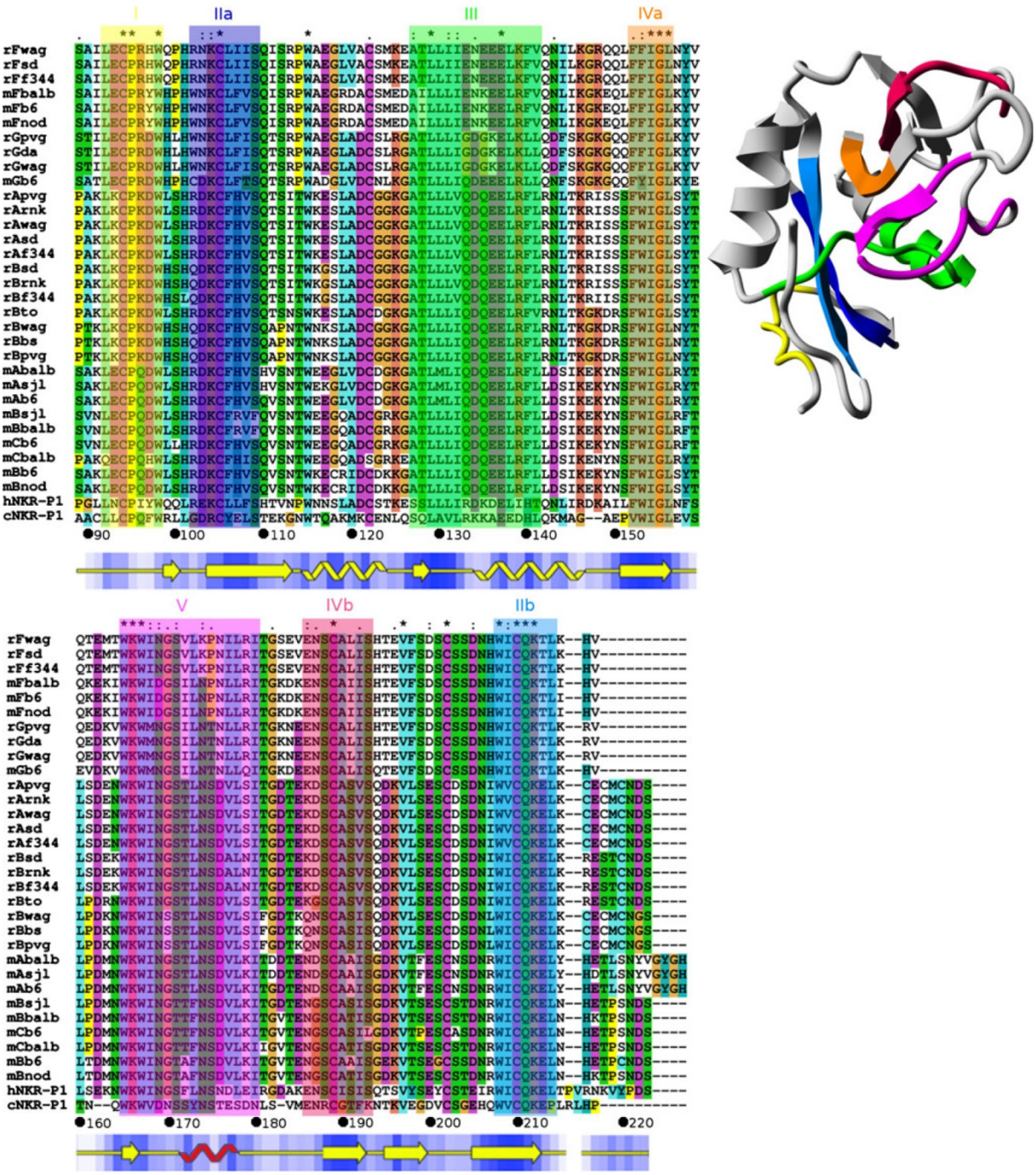


Fig. 12 Conserved regions among NKR-P1s and their positions in the structural model of rat NKR-P1A. A consensus secondary structure estimation is shown below the alignment in yellow. The red helix indicates that the helix is present in human NKR-P1 and rat NKR-P1B only. The color scheme underlying the secondary structure indicates buried (blue) and accessible (white) residues

structure that is variably present. The experimental data from vibrational thermal analysis would be consistent with the facile dynamics of this region observed in MD simulations, in which an underlying hydrophobic surface becomes exposed. Sequence and phylogenetic analysis suggests that this loop region evolved in concert with target specificity, adapting the chemical properties of the residues to specific target ligands while embedding them in a

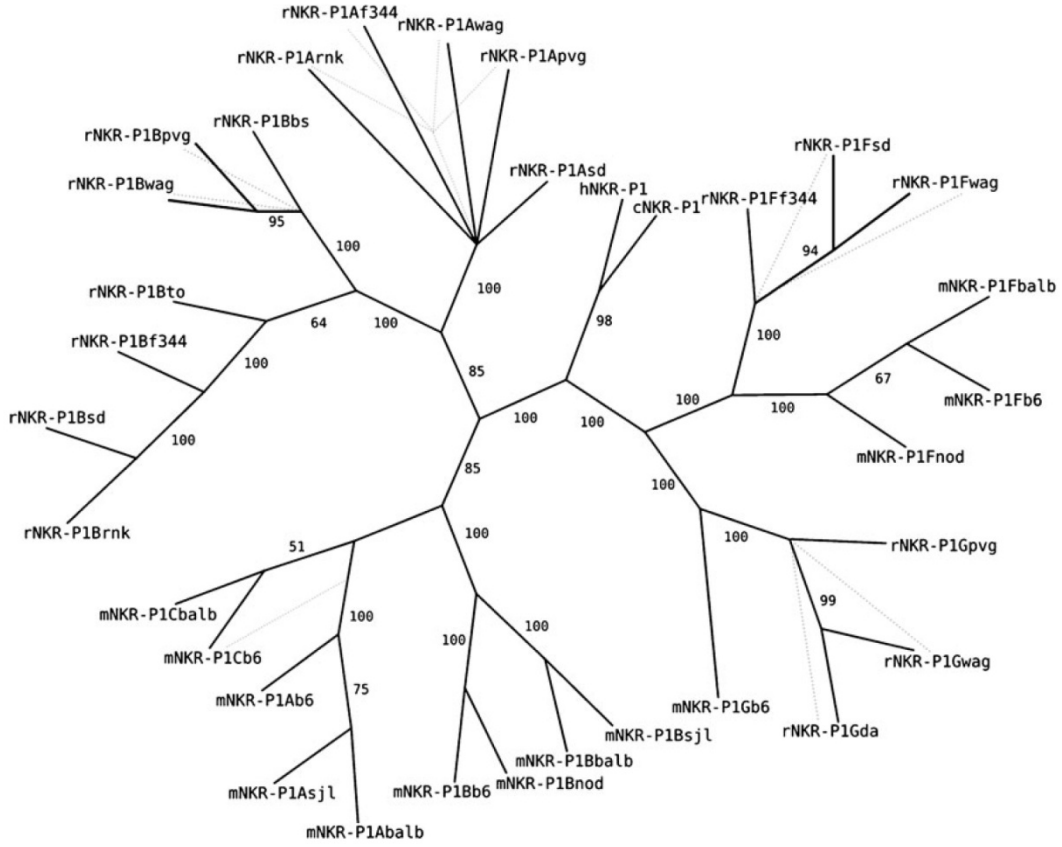


Fig. 13 A cladogram of all available NKR-P1s based on nucleotide sequences. If in some case different methods show either polytomy (full line) or further branching (dashed line), both are shown, however the polytomy is more relevant, the branching is only a possibility. Probability values are based on the maximum likelihood consensus tree and are displayed as percentages

structurally stable framework. The results thus suggest that the long loop plays a key role in ligand specificity, and that a conformational equilibrium is involved in ligand binding.

Homology modeling is based on the presumption that proteins with similar primary structures (> ~30% identity) have the same tertiary structure. Even though exceptions are known (e.g.,[63, 64]), the use of templates with functions close to those of the modeled targets in the present case suggested that this basic assumption should be correct, and that the core folds are likely to be similar. Local deviations can occur, and these may be detected by combining vibrational spectroscopy with computational modeling [65], as shown by the results presented here.

The resulting molecular models of NKR-P1 are available upon request.

Acknowledgments The Grant Agency of the Academy of Sciences of the Czech Republic and the Ministry of Education of the Czech Republic are gratefully acknowledged for support (LC 06010, No. KJB101120805, Nos. MSM6007665808, MSM 0021620835, MSM 0021620808, 1 M0505, AVOZ60870520, AV0Z50200510 respectively). Additionally, Z.S. was supported by the University of South Bohemia, grant GAJU 170/2010/P and O.V. by the E.C., project SPINE2-Complexes (031220). Access to the METACentrum computing facilities provided under the research intent MSM6383917201 is highly appreciated.

References

1. Lanier LL (1998) NK Cell receptors. Annu Rev Immunol 16:359–393
2. Vivier E, Tomasello E, Baratin M, Walzer T, Ugolini S (2008) Function of natural killer cells. Nat Immunol 9:503–510
3. Ljunggren HG, Karre K (1990) In search of 'missing self': MHC molecules and NK cell recognition. Immunol Today 11:237–244
4. Plíhal O, Byrtusová P, Pavlíček J, Mihók L, Ettrich R, Man P, Pompach P, Havlíček V, Hušáková L, Bezouška K (2004) The isoforms of rat natural killer cell receptor NKR-P1 display a distinct binding of complex saccharide ligands. Collect Czech Chem Commun 69:631–644
5. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248–254

6. Zhang D, Xie Y, Mrozek MF, Ortiz C, Davisson VJ, Ben-Amotz D (2003) Raman detection of proteomic analytes. Anal Chem 75:5703–5709
7. Kopecký V Jr, Baumruk V (2006) Structure of the ring in drop coating deposited proteins and its implication for Raman spectroscopy of biomolecules. Vibr Spectrosc 42:184–187
8. Dousseau F, Therrien M, Pézolet M (1989) On the spectral subtraction of water from FT-IR spectra of aqueous solutions of proteins. Appl Spectrosc 43:538–542
9. Voigt S, Mesci A, Ettinger J, Fine JH, Chen P, Chou W, Carlyle JR (2007) Cytomegalovirus Evasion of Innate Immunity by Subversion of the NKR-P1B:Clr-b Missing-Self Axis. Immunity 26:617–627
10. Giorda R, Rudert WA, Vavassori C, Chambers WH, Hiserodt JC, Trucco M (1990) NKR-P1, a signal transduction molecule on natural killer cells. Science 249(4974):1298–1300
11. Kveberg L, Dai KZ, Westgaard IH, Daws MR, Fossum S, Naper C, Vaage JT (2009) Two major groups of rat NKR-P1 receptors can be distinguished based on chromosomal localization, phylogenetic analysis and Clr ligand binding. Eur J Immunol 39:541–551
12. Dissen E, Ryan JC, Seaman WE, Fossum S (1996) An autosomal dominant locus, Nka, mapping to the Ly-49 region of a rat natural killer (NK) gene complex, controls NK cell lysis of allogeneic lymphocytes. J Exp Med 183:2197–2207
13. Appasamy PM, Kenniston TW, Brissette-Storkus CS, Chambers WH (1996) NKR-P1dim/TCR alpha beta + T cells and natural killer cells share expression of NKR-P1A and NKR-P1D. Nat Immunol 15(5):259–268
14. Giorda R, Trucco M (1991) Mouse NKR-P1 A family of genes selectively coexpressed in adherent lymphokine-activated killer cells. J Immunol 147(5):1701–1708
15. Carlyle JR, Mesci A, Ljutic B, Belanger S, Tai LH, Rousselle E, Troke AD, Proteau MF, Makrigiannis AP (2006) Molecular and genetic basis for strain-dependent NK1.1 alloreactivity of mouse NK cells. J Immunol 176(12):7511–7524
16. Kung SK, Su RC, Shannon J, Miller RG (1999) The NKR-P1B gene product is an inhibitory receptor on SJL/J NK cells. J Immunol 162(10):5876–5887
17. Carninci P, Hayashizaki Y (1999) High-efficiency full-length cDNA cloning. Methods Enzymol 303:19–44
18. Plougastel B, Matsumoto K, Dubbelde C, Yokoyama WM (2001) Analysis of a 1-Mb BAC contig overlapping the mouse Nkrp1 cluster of genes: cloning of three new Nkrp1 members, Nkrp1d, Nkrp1e, and Nkrp1f. Immunogenetics 53 (7):592–598
19. Lanier LL, Chang C, Phillips JH (1994) Human NKR-P1A A disulfide-linked homodimer of the C-type lectin superfamily expressed by a subset of NK and T lymphocytes. J Immunol 153(6):2417–2428
20. Kaufman J, Milne S, Gobel TW, Walker BA, Jacob JP, Auffray C, Zoorob R, Beck S (1999) The chicken B locus is a minimal essential major histocompatibility complex. Nature 401 (6756):923–925
21. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23:2947–2948
22. Felsenstein J (1989) PHYLIP – phylogeny inference package (version 3.2). Cladistics 5:164–166
23. Saitou N, Nei M (1987) The neighbor-joining method: A new method for reconstruction phylogenetics trees. Mol Biol Evol 4:406–425
24. Eck RV, Dayhoff MO (1966) Atlas of protein sequence and structure. National Biomedical Research Foundation, Silver Spring, MD
25. Felsenstein J (1973) Maximum-likelihood and minimum-steps methods for estimating evolutionary trees from data on discrete characters. Syst Zool 22:240–249
26. Fitch WM, Margoliash E (1967) Construction of phylogenetics trees. Science 155:279–284
27. Felsenstein J (1985) Confidence-limits on phylogenies – an approach using bootstrap. Evolution 39(4):783–791
28. Efron B (1982) Maximum-likelihood and decision-theory. Ann Stat 10(2):340–356
29. Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574
30. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410
31. Park H, Adsit FG, Boyington JC (2005) The 1.4 Å crystal structure of the human oxidized low density lipoprotein receptor lox-1. J Biol Chem 280:13593–13599
32. Snyder GA, Colonna M, Sun PD (2005) The structure of DC-SIGNR with a portion of its repeat domain lends insights to modeling of the receptor tetramer. J Biol Chem 347:979–989
33. Kolenko P, Skalova T, Vanek O, Stepankova A, Duskova J, Hasek J, Bezouska K, Dohnalek J (2009) The high-resolution structure of the extracellular domain of human CD69 using a novel polymer. Acta Crystallogr F 65:1258–1260
34. Llera AS, Viedma F, Sanchez-Madrid F, Tormo J (2001) Crystal structure of the C-type lectin-like domain from the human hematopoietic cell receptor Cd69. J Biol Chem 276(10):7312–7319
35. Brown J, O'Callaghan CA, Marshall ASJ, Gilbert RJC, Siebold C, Gordon S, Brown GD, Jones EY (2007) Structure of the fungal beta-glucan-binding immune receptor dectin-1: Implications for function. Protein Sci 16:1042–1052
36. Deng L, Cho S, Malchiodi EL, Kerzic MC, Dam J, Mariuzza RA (2008) Molecular architecture of the major histocompatibility complex class I-binding site of Ly49 natural killer cell receptors. J Biol Chem 283:16840–16849
37. Eswar N, Marti-Renom MA, Webb B, Madhusudhan MS, Eramian D, Shen M, Pieper U, Sali A (2006) Comparative protein structure modeling with MODELLER. Current Protocols in Bioinformatics, Wiley, Supplement 15, 5.6.1–5.6.30
38. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 26:283–291
39. Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis 18:2714–2723
40. Kabsch W, Sander C (1983) Dictionary of protein secondary structure: Pattern recognition of hydrogen bonded and geometrical features. Biopolymers 22:2577–2637
41. van der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ (2005) GROMACS: fast, flexible, and free. J Comput Chem 26(16):1701–1718
42. van Buuren AR, Marrink SJ, Berendsen HJC (1993) A molecular dynamics study of the decane/water interface. J Phys Chem 97:9206–9212
43. Mark AE, van Helden SP, Smith PE, Janssen LHM, van Gusteren WF (1994) Convergence properties of free energy calculations: α-cyclodextrin complexes as a case study. J Am Chem Soc 116:6293–6302
44. Tensmeyer LG, Kaufman EW (1996) Protein structure as revealed by nonresonance Raman spectroscopy. In: Havel HA (ed) Spectroscopic methods for determining protein structure in solutions. VCH, New York, pp 69–93
45. Overman SA, Thomas GJ Jr (1999) Raman markers of nonaromatic side chains in an α-helix assembly: Ala, Asp, Glu, Gly, Ile,

Leu, Lys, Ser, and Val residues of phage *fd* subunits. Biochemistry 38:4018–4027
46. Tuma R (2005) Raman spectroscopy of proteins: from peptides to large assemblies. J Raman Spectrosc 36:307–319
47. Miura T, Thomas GJ Jr (1995) Raman spectroscopy of proteins and their assemblies. In: Biswas BB, Roy S (eds) Subcellular Biochemistry, vol 24, Proteins: Structure, Function, and Engineering. Plenum, New York, pp 55–99
48. Miura T, Takeuchi H, Harada I (1989) Tryptophan Raman bands sensitive to hydrogen bonding and side-chain conformation. J Raman Spectrosc 20:667–671
49. Williams RW (1986) The secondary structure analysis using Raman amide I and amide III spectra. Method Enzymol 130:311–331
50. Dousseau F, Pézolet M (1990) Determination of the secondary structure content of proteins in aqueous solutions from their amide I and amide II infrared bands Comparison between classical and partial least-squares methods. Biochemistry 29:8771–8779
51. Baumruk V, Pancoska P, Keiderling TA (1996) Predictions of secondary structure using statistical analyses of electronic and vibrational circular dichroism and Fourier transform infrared spectra of proteins in H_2O. J Mol Biol 259:774–791
52. Fabian H, Mäntele W (2002) Infrared spectroscopy of proteins. In: Chalmers JM, Griffiths PR (eds) Handbook of vibrational spectroscopy, vol 5. Wiley, Chichester, pp 3399–3425
53. Barth A, Zscherp C (2002) What vibrations tell us about proteins. Q Rev Biophys 35:369–430
54. Barth A (2000) The infrared absorption of amino acid side chains. Prog Biophys Mol Biol 73:1–49
55. Miura T, Takeuchi H, Harada I (1988) Characterization of individual tryptophan side chains in proteins using Raman spectroscopy and hydrogen–deuterium exchange kinetics. Biochemistry 27:88–94
56. Siamwiza MN, Lord RC, Chen MC (1975) Interpretation of the doublet at 850 and 830 cm^{-1} in the Raman spectra of tyrosyl residues in proteins and certain model compounds. Biochemistry 14:4870–4876
57. Noda I (1990) Two-dimensional infrared (2D IR) spectroscopy: Theory and applications. Appl Spectrosc 44:550–561
58. Noda I (1993) Generalized two-dimensional correlation method applicable to infrared, Raman, and other types of spectroscopy. Appl Spectrosc 47:1329–1336
59. Noda I, Ozaki Y (2004) Two-dimensional correlation spectroscopy: Applications in vibrational and optical spectroscopy. Wiley, Chichester
60. Miura T, Takeuchi H, Harada I (1991) Raman spectroscopic characterization of tryptophan side chains in lysozyme bound to inhibitors: Role of the hydrophobic box in the enzimatic function. Biochemistry 30:6074–6080
61. Weis WI, Taylor ME, Drickamer K (1998) The C-type lectin like superfamily in the immune system. Immunol Rev 163:19–34
62. Pompach P, Man P, Kavan D et al (2009) Modified electrophoretic and digestion conditions allow a simplified mass spectrometric evaluation of disulfide bonds. J Mass Spectrom 44:1571–1578
63. Kabsch W, Sander C (1985) Identical pentapeptides with different backbones. Nature 317:19–25
64. Sternberg MJE, Islam SA (1990) Local protein sequence similarity does not imply a structural relationship. Protein Eng 4:125–131
65. Kopecký V Jr, Ettrich R, Hofbauerová K, Baumruk V (2004) Vibrational spectroscopy and computer modeling of proteins: solving structure of α_1-acid glycoprotein. Spectrosc Int J 18:323–330

PAPER III

Kolenko, P., Rozbesky, D., Vanek, O., Kopecky, V., Jr., Hofbauerova, K., Novak, P., Pompach, P., Hasek, J., Skalova, T., Bezouska, K., Dohnalek, J.

Molecular architecture of mouse activating NKR-P1 receptors.

J. Struct. Biol. **175**, 434-41 (2011)

My contribution to the publication: *research performing (protein expression, refolding in vitro, purification, disulfide mapping, chemical cross-linking), data collection, data analysis and interpretation, manuscript writing*

Journal of Structural Biology 175 (2011) 434–441

Contents lists available at ScienceDirect

Journal of Structural Biology

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

Molecular architecture of mouse activating NKR-P1 receptors

Petr Kolenko [a,1], Daniel Rozbeský [b,c,1], Ondřej Vaněk [b,c], Vladimír Kopecký Jr. [d], Kateřina Hofbauerová [c,d], Petr Novák [c], Petr Pompach [c], Jindřich Hašek [a], Tereza Skálová [a], Karel Bezouška [b,c], Jan Dohnálek [a,*]

[a] Institute of Macromolecular Chemistry, v.v.i., Academy of Sciences of the Czech Republic, Heyrovského náměstí 2, 16206 Prague 6, Czech Republic
[b] Department of Biochemistry, Faculty of Science, Charles University in Prague, Hlavova 8, 12840 Prague 2, Czech Republic
[c] Institute of Microbiology, v.v.i., Academy of Sciences of the Czech Republic, Vídeňská 1083, 14220 Prague 4, Czech Republic
[d] Institute of Physics, Faculty of Mathematics and Physics, Charles University in Prague, Ke Karlovu 5, 12116 Prague 2, Czech Republic

ARTICLE INFO

Article history:
Received 9 March 2011
Received in revised form 28 April 2011
Accepted 5 May 2011
Available online 12 May 2011

Keywords:
NK cell
C-type lectin-like receptor
X-ray structure
Receptor–ligand interaction

ABSTRACT

Receptors belonging to NKR-P1 family and their specific Clr ligands form an alternative missing self recognition system critical in immunity against tumors and viruses, elimination of tumor cells subjected to genotoxic stress, activation of T cell dependent immune response, and hypertension. The three-dimensional structure of the extracellular domain of the mouse natural killer (NK) cell receptor mNKR-P1Aex has been determined by X-ray diffraction. The core of the C-type lectin domain (CTLD) is homologous to the other CTLD receptors whereas one quarter of the domain forms an extended loop interacting tightly with a neighboring loop in the crystal. This domain swapping mechanism results in a compact interaction interface. A second dimerization interface resembles the known arrangement of other CTLD NK receptors. A functional dimeric form of the receptor is suggested, with the loop, evolutionarily conserved within this family, proposed to participate in interactions with ligands.

1. Introduction

Natural killer (NK) cells are large granular lymphocytes able to recognize and lyse a broad range of target cells including tumors, virus infected, antibody coated, and stressed cells (Vivier et al., 2008). NK cells eliminate these targets through cellular cytotoxicity mechanisms initiated through a direct cell–cell contact. According to the "missing self" hypothesis NK cells check for the presence of established cell surface molecules, such as MHC class I expressed on normal cells under physiological conditions, and initiate the lysis of those cells in which these markers of "good health" are missing (Kärre, 2008). Since MHC class I glycoproteins have evolved rapidly in mammals, two families of specific receptors recognizing these structures have emerged during the mammalian evolution in the form of both immunoglobulin (e.g. KIR) and C-type lectin (e.g. Ly49) inhibitory receptors delivering dominant negative signals to NK cells (Ryan and Seaman, 1997; Valiante et al., 1997). Structural biology of these interactions is understood in considerable detail since the crystal structures of some MHC class I-KIR as well as MHC class I-Ly49 complexes are available (Boyington et al., 2000; Deng et al., 2008). Nevertheless, an alternative molecular mechanism for the "missing self" recognition of tumor cells was later discovered based on the interaction of the inhibitory NKR-P1 receptors (CD161) with CLEC2D gene family members such as C-type lectin related (Clr, also known as Ocil) or Lectin-like transcript-1 (LLT1) ligands in rodents or human, respectively (Iizuka et al., 2003; Carlyle et al., 2004; Aldemir et al., 2005; Rosen et al., 2005). This recognition is based on lectin–lectin interaction in which the C-type lectin domains (CTLDs) interact apparently on the basis of protein–protein recognition (Carlyle et al., 2004). This type of recognition may be relevant not only for inhibitory NKR-P1 receptors (mouse NKR-P1B and its allelic variants – some of them also termed as NKR-P1D; rat NKR-P1B and G; human NKR-P1), but also for their activating isoforms (mouse NKR-P1A, C and F; rat NKR-P1A and F) (Mesci et al., 2006; Aust et al., 2009; Kveberg et al., 2009). For example, the prototype receptor of the family, rat NKR-P1A (Giorda et al., 1990), has been recently shown to interact with Clr11 (a functional orthologue of mClrb) in rat belonging to the PVG.7B (PVG-RT7^b) strain (Kveberg et al., 2009). However, molecular details of these interactions remain unknown.

Abbreviations: Clr, C-type lectin receptor; CTLD, C-type lectin like domain; DLS, dynamic light scattering; DSC, differential scanning calorimetry; DTT, dithiothreitol; EDC, 1-ethyl-3-(3′-dimethylaminopropyl) carbodiimide; FT-MS, Fourier-transform mass spectrometry; ITAM, immunoreceptor tyrosine-based activation motif; KIR, killer-cell immunoglobulin-like receptor; LC-ESI–FTICR-MS, liquid chromatography coupled with electrospray ionization and Fourier-transform ion-cyclotron-resonance mass spectrometry; MMR4, domain 4 of the macrophage mannose receptor; MR, molecular replacement; MS, mass spectrometry; NKR, natural killer cell receptor; PVDF, polyvinylidene fluoride; Rmsd, root mean square deviation; TCEP, tris(2-carboxyethyl)phosphine; VDW, van der Waals.

* Corresponding author. Fax: +420 296809410.
E-mail address: dohnalek007@gmail.com (J. Dohnálek).
[1] These authors have contributed equally to the manuscript.

1047-8477/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jsb.2011.05.001

Functionally, cross-linking of rat or mouse activating NKR-P1 receptors by antibodies results in rapid cellular activation connected with generation of inositol phosphates, and an increase in intracellular calcium levels (Ryan et al., 1991), enhanced lysis of certain tumor cells (Ryan et al., 1995), and production of cytokines by the stimulated NK cell such as interferon-γ (Arase et al., 1996). Activating NKR-P1 receptors signal through their association with FcRγ chain (Arase et al., 1997) that is followed in 5 min by phosphorylation of their ITAM motifs by tyrosine kinase $p56^{lck}$, and subsequent downstream signaling propagated by Syk kinase (Cerny et al., 1997; Ljutic et al., 2005). In mice, studies of the molecular mechanisms of signaling by these receptors have been complicated by a lack of monoclonal antibodies specific for individual NKR-P1 isoforms as well as by genetic and transcriptional heterogeneity among these receptors in individual experimental mouse strains. Notably, it has been shown by a detailed genetic study that mouse NKR-P1C expressed in C57BL/6 mice (NKR-P1C(B6)) has a total of 25 amino acid substitutions compared to NKR-P1C found in BALB/c mice (NKR-P1C(BALB)) including the mutation of Ser^{191} to Thr^{191} inactivating the reactivity with the established anti-NK1.1 antibody PK136 (Carlyle et al., 2006). Also, NKR-P1C(B6) occurs in three transcriptional isoforms differing in the neck region (Giorda and Trucco, 1991). However, several recent studies succeeded in generation of specific monoclonal antibodies and/or establishing functional assays making it possible to probe the details of expression and function of individual NKR-P1 and Clr isoforms (Kveberg et al., 2006, 2009; Aust et al., 2009). Two major groups of NKR-P1 receptors (A–D and F–G) can be distinguished in both mouse and rat based on chromosomal localization, phylogenetic analysis and Clr ligand binding, which demonstrate conservation of Clr ligand-binding properties independent of NKR-P1 signaling function. Moreover, inhibitory NKR-P1B/D seems to distinguish two functionally distinct populations of NK cells differing in Ly49 and CD94/NKG2 receptor expression pattern with NKR-P1B/D^+ cells showing higher cytotoxic potential than NKR-P1B/D^- cells, a finding reminiscent of inhibitory Ly49-based licensing of NK cells (Kim et al., 2005). The importance of missing self recognition by the inhibitory NKR-P1B/D–Clrb receptor pair was recently underlined by a study of chemotherapy-induced tumor genotoxic stress promoting sensitivity to NK cell cytotoxicity by downregulation of Clrb (Fine et al., 2010). On the other hand, this missing self axis could be also subverted in viral evasion mechanisms (Voigt et al., 2007). Beside its role in missing self recognition, NKR-P1–Clr receptor system is involved in T-cell activation or enhanced NK cell lysis of tumors following their cross-talk with dendritic cells (Tian et al., 2005a,b; Yang et al., 2006), and in hypertension (Taherzadech et al., 2010). Human NKR-P1 recognition of LLT1, a functional orthologue of rodent Clrb, is again suggested to be involved in regulation of the cross-talk between NK cells and antigen presenting cells (Rosen et al., 2008). Importantly, it has been shown that LLT1 is upregulated on glioblastoma cancer cells, one of the most lethal tumors, and its inhibitory interaction with NKR-P1 expressed on NK cells contributes to its immunosuppressive properties (Roth et al., 2007). Apart from its role on NK cells, human NKR-P1 plays also an important role on T cells, especially on a Th17 cells subtype with distinct tissue homing and functional properties connected with intestinal inflammation, chronic hepatitis, arthritis and potentially with other infectious and inflammatory diseases of man (Kleinschek et al., 2009; Billerbeck et al., 2010).

No three dimensional structure of any of the NKR-P1 family of receptors or their Clr ligand has been published so far. Our recent work analyzing the entire NKR-P1 family revealed two loop regions that do not correspond to the conserved structural elements of the C-type lectin family (Sovova et al., 2010). One of these regions displays high flexibility but is anchored by conserved sequences suggesting that its relative position to the rest of the domain might be variable. NKR-P1 receptors are expected to interact with their ligands via their CTL domains. They have a characteristic loop-in-loop fold stabilized by at least two conserved disulfides, and a network of hydrophobic and polar interactions (Zelensky and Gready, 2005). The second, surface loop of the domain is usually involved in ligand binding and this part of the chain was also observed to participate in domain swapping leading to formation of stable multimers in crystal structures of some CTLD proteins (Mizuno et al., 1997, 2001; Liu and Eisenberg, 2002; Maita et al., 2003). We describe here the first three-dimensional structure of the soluble extracellular ligand-binding domain of mouse NKR-P1A (NKR-P1Aex). The evolutionarily conserved loop of this receptor that forms a notable structural feature is involved in domain swapping, and is suggested to play a role in interactions with Clr ligands.

2. Materials and methods

2.1. Recombinant proteins

Total RNA was isolated from spleens of C57BL/6 mice with RNA Blue (Top-Bio), and used as a template for single stranded cDNA synthesis (Superscript III reverse transcriptase, Invitrogen). DNA fragment coding for the extracellular part of mouse NKR-P1A (NKR-P1Aex, Ser^{89}–His^{227}) was amplified by PCR using primers incorporating *Nde*I and *Hind*III sites and subcloned into pET-30a expression vector. Protein was expressed in *Escherichia coli* BL-21 (DE3) Gold (Stratagene) and isolated as inclusion bodies. These were solubilized in 6 M guanidine–HCl (pH 8.0) with 10 mM DTT, and the protein was refolded by rapid dilution into a 100-fold excess of the refolding buffer (50 mM Tris–HCl (pH 8.5), 0.4 M L-arginine, 100 mM $CaCl_2$, 9 mM cysteamine, 3 mM cystamin, and 1 mM PMSF). The refolding mixtures were dialyzed twice against 8 l of 15 mM Tris–HCl (pH 8.5), 9 mM NaCl, and 1 mM NaN_3. The refolded protein was purified using ion exchange chromatography on Q-Sepharose FF (1.0 × 6.0 cm), followed by gel filtration on a Superdex 75 HR 10/30 column (both GE Healthcare) eluted with 10 mM HEPES pH 7.4 with 150 mM NaCl and 1 mM NaN_3. Purified protein were concentrated to 10 mg/ml, and stored at 4 °C.

2.2. Protein characterization

Protein identity was confirmed by MS peptide mapping after *in gel* digestion (Pompach et al., 2009), and *N*-terminal sequencing (Procise 491, Applied BioSystems) using BioBrene-precycled Protein Filters (Applied BioSystems), or PVDF blots. Correct disulfide bond arrangement of all recombinant proteins was determined by digestion with trypsin or Asp–N, and subsequent FT-MS analysis of disulfide linked peptides (Pompach et al., 2009). Size distribution and polydispersity of NKR-P1Aex were assessed by DLS at concentration 1 mg/ml in 10 mM HEPES (pH 7.0), 50 mM NaCl, and 1 mM NaN_3 at 18 °C using a Zetasizer Nano (Malvern Instruments). Protein stability of mNKR-P1Aex was verified by DSC with a 6100 Nano II DSC (Microcalorimetry Sciences). Protein diluted to 0.2 mg/ml in 15 mM Tris–HCl (pH 8.0), 150 mM NaCl, and 1 mM NaN_3 was scanned from 5 to 90 °C at heating rate 1 °C min^{-1}. Sedimentation analysis was performed using a ProteomeLabXL-I analytical ultracentrifuge (BeckmanCoulter). NKR-P1Aex (0.2 mg/ml) was dialyzed against 10 mM HEPES (pH 7.0), 50 mM NaCl and 1 mM NaN_3 buffer, used also as a reference. Sedimentation velocity experiment was carried out at 48,000 rpm and 20 °C, absorbance scans were recorded at 280 nm in 5 min intervals with 30 μm spatial resolution. Buffer density and NKR-P1Aex partial specific volume were estimated in SEDNTERP 1.09 (www.jphilo.mailway.com). Data were analyzed with SEDFIT 12.1 (Schuck, 2000). The sedimentation equilibrium experiment was performed with

NKR-P1Aex concentration 0.1 mg/ml at 10–12–14–16–18–20–22–24–26–28–30–32,000 rpm at 4 °C. Absorbance data were collected at 280 nm by averaging 20 scans with 10 μm spatial resolution after 20 h of achieving equilibrium and were globally analyzed with SEDPHAT 8.2 (Schuck, 2003) using a non-interacting discrete species model.

2.3. Structure determination

NKR-P1Aex was crystallized by hanging-drop vapor-diffusion method with initial screening using Index and Crystal Screen (Hampton Research), and polymer screens PolyA and PolyB (Skalova et al., 2010). Optimized drops were prepared with 1 μl protein and 1 μl reservoir solution (0.3 M ammonium phosphate) at 291 K. The bipyramidal crystals, grown in 4 days, were cryoprotected with 30% (v/v) ethylene glycol. X-ray diffraction data were collected from a single crystal at BESSY II (beam line BL14.1) at temperature 100 K with a MARMOSAIC 225 CCD detector. The data were processed with MOSFLM (Collaborative Computational Project, Number 4, 1994) and SCALA (Evans, 2005) up to 1.7 Å resolution in space group $I4_122$ (Table 1). The phase problem was solved by molecular replacement using BALBES (Long et al., 2008), a modified extracellular domain of hCD69 was used as a search model (PDB ID 3HUP). The model was refined with REFMAC5 (Murshudov et al., 1997), and manually corrected with COOT (Emsley and Cowtan, 2004). Refinement was monitored using the statistic R_{free}. The last refinement cycle was performed against all measured reflections. The model and data quality were assessed with MOLPROBITY (Lovell et al., 2003) and SFCHECK (Vaguine et al., 1999). 100% of residues are in the favored region of the Ramachandran plot. Molecular graphics was created with PyMOL (DeLano Scientific) and electrostatic potential distribution was calculated with APBS (Baker et al., 2001).

2.4. Raman spectroscopy

Raman spectra of NKR-P1Aex solution (15 mM Tris–HCl, pH 8.0, 150 mM NaCl, protein concentration 10.4 mg/ml) were recorded on an instrument based on Spex 270 M single spectrograph with 1800 grooves/mm grating (Jobin–Yvon), a holographic notch-plus filter (Kaiser Optical Systems), and a liquid nitrogen cooled CCD (Princeton Instruments). Samples in a 5 μl-capillary micro-cell were excited with a 532.2 nm NdYAG laser (300 mW). The spectra were integrated 3600 s. Spectra in microscopic measurements with an HR800 Raman microspectrometer (Horiba Jobin Yvon) were collected using 600 grooves/mm grating and a liquid nitrogen cooled CCD (Symphony). Protein crystals were measured directly in hanging drops (10 mM Tris–HCl, pH 8.0, 0.3 M ammonium phosphate, 150 mM NaCl, and 1 mM NaN_3) using He–Ne laser 632.82 nm (10 mW), and 50× microscope objective (N.A. 0.55, $f = 180$, Olympus). Spectra were integrated 600 s with resolution ca. 5 cm^{-1}. Wavenumber scales were calibrated with neon glow-lamp lines, thus Raman frequencies of well-resolved bands are accurate to ±0.5 cm^{-1}. Buffer spectra were subtracted and the background polynomial correction was applied.

Table 1
Summary of crystallographic data and structure refinement details.

X-ray source, beamline	BESSY, BL14.1
Wavelength (Å)	0.91841
No. oscillation images	100
Oscillation range (°)	1.0
Resolution range (Å)	15.00–1.70 (1.79–1.70)
No. observed reflections	121662 (12277)
No. unique reflections	16416 (2319)
Data completeness (%)	99.7(99.0)
Average redundancy	7.4 (5.3)
Space group	$I4_122$
Unit cell parameters	–
$a = b$ (Å), c (Å)	59.91, 159.30
R_{merge}	0.070 (0.704)
$\langle I \rangle / \langle \sigma I \rangle$	14.3 (2.1)
Wilson B (Å^2)	24
No. residues (No. non-H atoms)	124 (1018)
No. water molecules	111
No. all non-H atoms	1134
Rmsd bond lengths from ideal (Å)	0.019
Rmsd bond angles from ideal (°)	1.741
Average B-factors (Å^2)	
Protein atoms	29
Solvent molecules	39
All atoms	30
R_{work} (R_{free})	0.195 (0.239)
R	0.197
No. reflections for R_{free} calculation, (%)	829, (5%)

2.5. Chemical cross-linking and mass spectrometry

NKR-P1Aex (0.2 mg/ml) was cross-linked by 25× and 100× molar excess of EDC reagent in 50 mM pyridine–HCl (pH 6.5) and 50 mM NaCl for 4 h at room temperature. Reaction mixture was resolved on SDS–PAGE gel, a band of the cross-linked dimeric protein was excised, subjected to *in-gel* reduction by TCEP, alkylation by iodoacetamide, and digestion by trypsin and Asp–N. Digested peptides were extracted and subjected to a one-step cleanup (Micro Traps, Michrom Bioresources, Inc.). Mass of the generated peptides was measured using LC-ESI–FTICR-MS. Data processing and interpretation were performed using the Data Analysis software 4.0 (Bruker Daltonics) and MS2Links (http://ms3d.org).

2.6. Accession numbers

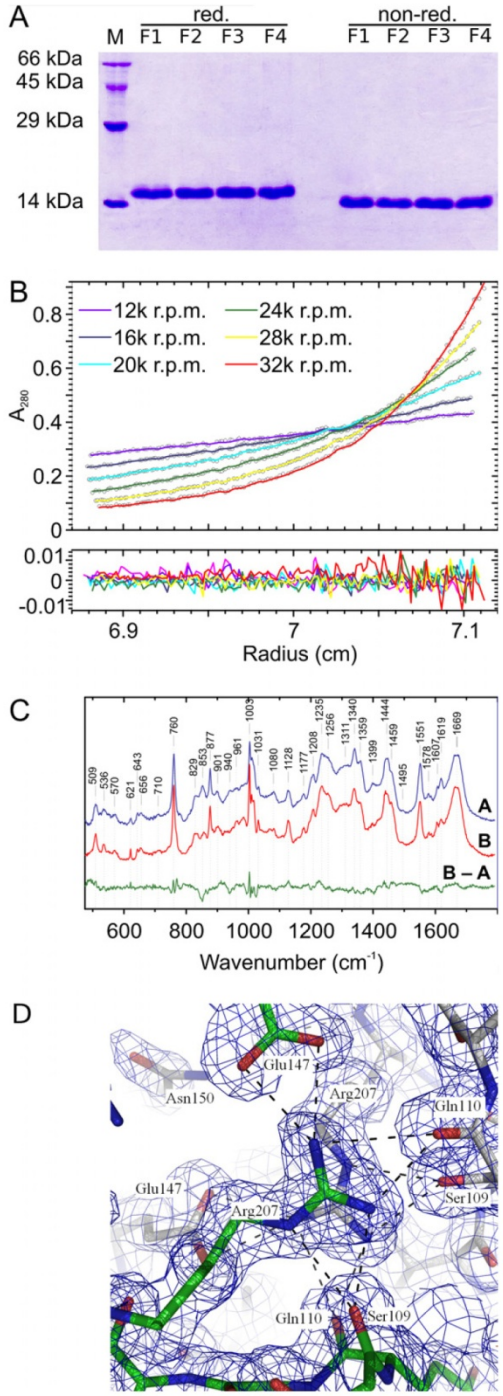
The coordinates and structure factors for NKR-P1Aex have been deposited in the Protein Data Bank, PDB ID 3M9Z.

3. Results

3.1. Structure of NKR-P1Aex

We developed methods for expression and purification of a recombinant soluble form of mouse NKR-P1A (residues Ser^{89}–His^{227}) encompassing one half of the stalk region, and the extracellular ligand binding domain (NKR-P1Aex). In size exclusion chromatography, NKR-P1Aex eluted as one symmetric peak at a position corresponding to a monomer. When analyzed by SDS–PAGE under reducing and non-reducing conditions, NKR-P1Aex displayed a shift in mobility under non-reducing conditions due to a more compact arrangement of the protein cross-linked by three disulfide bridges (Fig. 1A). An analysis of sedimentation equilibrium (Fig. 1B) further confirmed the monomeric state of the protein: the calculated molecular weight of 15,976 ± 101 Da correlated well with the theoretical molecular mass (15,986 Da). On the other hand, a sedimentation velocity experiment revealed a tiny fraction of dimeric species (Fig. A1). The identity, monodispersity, and high purity of protein preparation was confirmed by high resolution mass spectrometry. Dynamic light scattering (DLS) experiments yielded a particle diameter of 4 nm corresponding to an expected size of one monomer and also to sedimentation velocity data (sedimentation coefficient $s_{20,w} = 1.8$ S). NKR-P1Aex is stable at least up to 55 °C during differential scanning calorimetry (DSC) measurement, reaching its melting temperature at 69 °C when it quickly undergoes irreversible denaturation, ranking this protein as a stable one. The highly purified protein crystallized using the hanging drop vapor diffusion method providing tetragonal crystals.

The structure of NKR-P1Aex was determined by molecular replacement using an extracellular domain of human CD69 (hCD69) with 30% sequence identity (see Fig. 2I) as a search model (Vanek et al., 2008; Kolenko et al., 2009). The structure was refined to 1.7 Å resolution and a very good agreement with the experimental data (Table 1 and Fig. 1D, PDB ID 3M9Z). It contains one protein chain in asymmetric unit (residues Leu^{92}–Tyr^{215}) with a phosphate ion bound close to an interface with a symmetry related molecule. The overall fold of NKR-P1Aex is similar to that of hCD69 but differs significantly in a thirty residue region Tyr^{158}–Asp^{187}, which exhibits domain swapping with a symmetry related molecule. The N- and C-termini are located close to each other on surface of the molecule and opposite the extended swapped region (Fig. 2A). The secondary structure of the extracellular domain is formed by three alpha helices (α1–3), and two antiparallel beta sheets (strands β1, 2, 6, and β3–5). Three disulfide bridges stabilize the overall structure of the compact core region. This region (residues Leu^{92}–Leu^{160} and Ser^{188}–Tyr^{215}) shares high structural similarity with hCD69, the root-mean-square deviation between equivalent C_α atoms of the corresponding secondary structure elements of the extracellular domain of hCD69 and the compact core region of NKR-P1Aex is 1.57 Å (secondary structure matching, Krissinel and Henrick, 2004). NKR-P1Aex is found in the crystal interacting with several tightly packed neighboring molecules. Surface area buried in protein–protein interactions is one of the criteria helpful in recognition of biologically relevant complexes (Krissinel and Henrick, 2007). One molecule of NKR-P1Aex interacts with a total of seven neighboring molecules in the crystal. Only two of the contacts involve sufficiently large areas of protein surface and are of a potential interest for true complex formation of this receptor.



3.2. Intermolecular interactions observed in the crystal

The most extensive interaction interface (type A, Fig. 2B) is the region of the swapped loop with an interaction surface area of about 1700 Å^2 involving 25 hydrogen bonds (eight main chain–main chain). The loop binds to the surface of another molecule so that it mimics the situation in hCD69 where a part of the protein chain corresponding to the extended loop is placed back onto the compact core (Fig. 2H). A sequence identity of about 30% between the swapped region in NKR-P1Aex, and the corresponding region of hCD69 is roughly the same as for the entire extracellular domain. Therefore, this structural behavior in NKR-P1Aex would have been quite difficult to predict on the basis of sequence alignment. The protein dimer formed by these tightly interacting swapped loops results in a structure composed of two globules with a deep narrow groove between them. The groove is lined by Tyr^{149} in the middle, and Glu^{200} and Lys^{196} at the bottom. The size of the groove (20 Å long, 8 Å wide and 14 Å deep) would allow binding of small ligands (Fig. 2B).

The second largest interaction area of about 650 Å^2 (type B, Fig. 2C) is found in the region close to the bound phosphate ions, formed mostly by interactions of residues 110–112 in contact with the tip of helix α2 (residues 145–147) and of Arg^{207} of strand β6 in a network of H-bonds between the two regions. At the same time, Arg^{207} participates in a central π–π stacking interaction with the opposite Arg$^{207'}$ of the neighbor (Fig. 1D). This interface involves eight H-bonds (two main chain–main chain). A dimer constructed

Fig.1. Purification and characterization of the soluble form of mouse NKR-P1A. (A) Homogeneity verification of the size-exclusion chromatography fractions (F1–F4) by SDS–PAGE under reducing and non-reducing conditions; lane M represents molecular weight markers. (B) Equilibrium sedimentation distribution of NKR-P1Aex. The upper panel shows absorbance data (circles, every second scan is shown) for the respective rpm values with fitted curves (non-interacting discrete species model, lines), lower panel shows residuals derived from the fitted data. (C) Raman spectra of NKR-P1Aex measured in (A) solution and (B) crystal. The spectral difference (B – A) is shown in the lower part. See also Tables A1 and A2. (D) Example of agreement between model and electron density at the interface of dimer B with stacked arginines in the center; protein as sticks, color coded by atom types, chain A carbon green, chain B grey, 2Fo–Fc Fourier contoured in blue at level 1σ, hydrogen bonds as dashed lines, the interface residues are labeled.

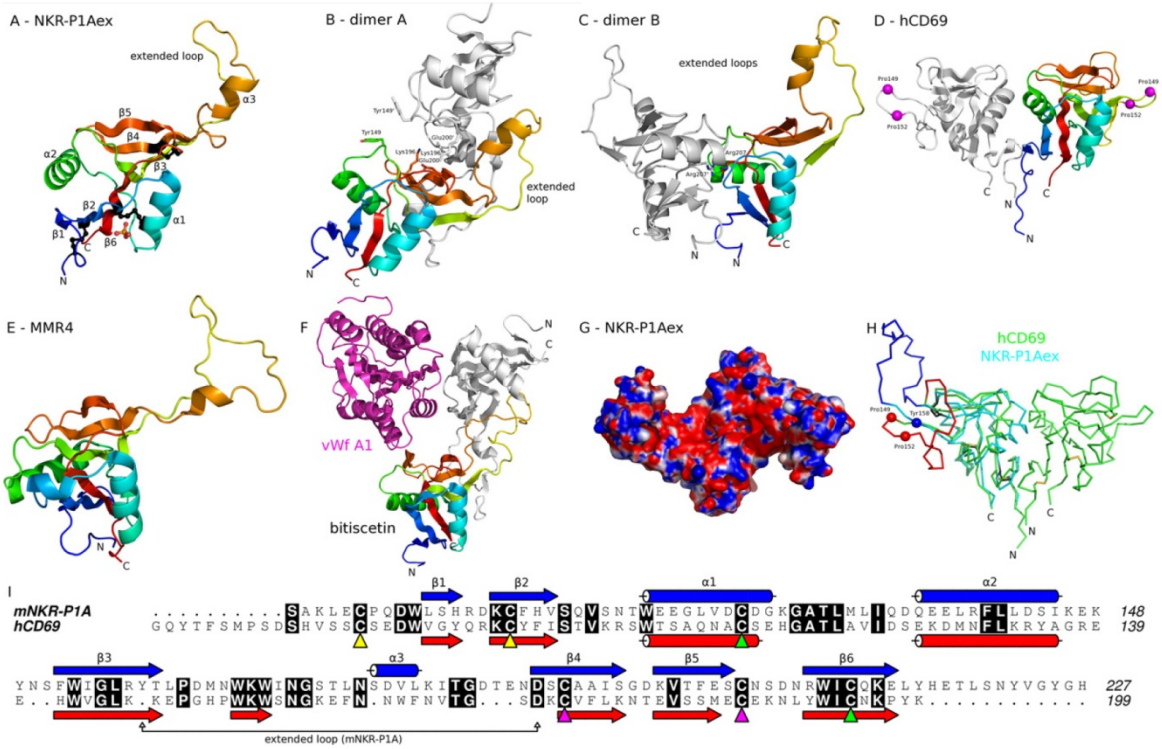


Fig.2. Secondary structure representation of NKR-P1Aex and its oligomers in relation to the previously known C-type lectin-like proteins. (A) Structure of one chain of NKR-P1Aex (secondary structure elements are marked, color ramp from blue to red along the protein chain, disulfide bridges and phosphate ion as balls and sticks), the asymmetric unit of crystal. (B) Domain-swapping observed in the structure leading to dimer type A of NKR-P1Aex, coloring as in (A) with the second chain in grey, residues contributing to the cleft walls are marked. (C) Dimer of type B (CD69-like) of NKR-P1Aex observed in the structure, coloring as in (A) with the second chain in grey, Arg^{207} involved in the central interaction is marked. (D) Dimer of the extracellular domain of hCD69 used for the molecular replacement search, PDB ID 3HUP, coloring as in (A) with the second chain in grey, magenta spheres mark the discussed prolines of the surface loop. (E) Domain 4 of the macrophage mannose receptor (1EGG), also able to form a domain-swapped dimer, coloring as in (A). (F) Domain-swapped dimer of snake venom bitiscetin (chains colored by ramp from blue to red and in grey) binding to von Willebrand factor A1 domain (purple, 1UEX). (G) Electrostatic surface potential projected onto the surface of NKR-P1Aex type A dimer viewed from the top, potential levels −2kT red, +2kT blue. (H) Superposition of hCD69 dimer (green and red) and NKR-P1Aex monomer (cyan and blue) to illustrate the difference in folding of the extended loop, the red and blue spheres mark the positions of the discussed critical residues, disulfide bridges in yellow. (I) Structure based sequence alignment of NKR-P1Aex and of the extracellular domain of hCD69; secondary structure elements are indicated by color graphics and the extended loop of the mouse receptor by solid line, colored triangles indicate cysteins involved in disulfide bridges. See also Fig. A2 for sequence comparison to other members of the NKR-P1 family.

by this interaction alone resembles to some extent that of hCD69 (Fig. 2D, Vanek et al., 2008; Kolenko et al., 2009) but comprises a different mutual orientation of monomers, and a different set of interactions.

Several amino acid residues of NKR-P1Aex can be identified as crucial for the extended conformation of the loop, and for its interactions with surrounding protein: Leu^{119}, Gln^{135}, Tyr^{158}, Met^{163}, Trp^{165}, Trp^{167}, Ile^{168}, Asn^{174}, Asp^{176}, Ile^{180}, Asp^{183} and Phe^{199}. Some of the corresponding residues are utilized by hCD69 in formation of the compact CTL domain, and so a unique set of the most important residues involved in the actual domain swapping is reduced to Tyr^{158}, Met^{163} and Phe^{199}. On the other hand, the compact fold of hCD69 in this region seems to rely on the presence of the PGXP (149–152) and WF (164–165) motives. NKR-P1Aex lacks the double-proline motif that serves human CD69 to flip the loop to form the compact globular domain (Fig. 2I), packing the chain against the core and not allowing the exposure to solvent. The mechanism of binding to the opposite chain is very similar to that of binding of the equivalent loop of hCD69 to the same chain. Judged by the structural differences, residues Leu^{119}, Phe^{199}, Tyr^{158}, Leu^{160} and Met^{163} disallow formation of a compact monomer of the hCD69 type and promote the inter-chain interaction. The typical sequence signatures of both types of packing can be distinguished and comparison of amino acid sequences of the NKR-P1 proteins (Fig. A2) and analysis of the structures suggest that Tyr^{158} has a particularly important role. Of the above named residues, Tyr^{158}/Phe^{158} is conserved throughout the NKR-P1 receptor family and Met^{163} in most cases. On the other hand, the PGXP and WF motives found in hCD69 are missing in the NKR-P1 family (Fig. A2).

If both types of interactions (A and B) are taken into account simultaneously, a tetramer of NKR-P1Aex can be obtained. Except for the above mentioned types of dimeric interfaces, Tyr^{149} participates in VDW contacts with the same residue of an opposite protein chain of the tetramer. Such state, though, would still enable binding of small ligands into the clefts described for dimer A.

The Raman spectra of NKR-P1Aex in solution and in crystal are almost identical (Fig. 1C), and only subtle structural changes could be observed upon the detailed inspection of the spectra (see the assignments of the Raman bands in Table A1). Aromatic amino acid residues display higher values of torsion angles as the sample density and orderliness increase, which is in agreement with the slightly higher average torsion angle determined for the crystal structure. The changes in the region of νS–S vibrations indicate that cysteines in the crystal have more precisely defined conformations than in solution. Conformations of S–S bridges – two in *gauche–gauche–gauche* and one close to *trans-gauche-trans*

conformation – are in an excellent agreement with the crystal structure, and are also preserved in solution. Raman spectroscopy methods for secondary structure determination overestimated the α-helical and β-sheet content. However, the determined secondary structure is still in a good agreement with the crystal structure. The comparison of an amide I region (1630–1700 cm^{-1}) indicates that the secondary structure elements occurring in the crystal structure are nearly identical to those present in solution (Table A2).

3.3. Dimer B of mouse NKR-P1Aex observed by chemical cross-linking

As described above, the stronger interaction of type A occurs through domain swapping in the crystal. On the other hand, while NKR-P1Aex was found to be almost exclusively monomeric, the natural arrangement of most NK receptors with CTLD at the surface of immune cells is in covalent homodimers (Mesci et al., 2006). The interaction of type B, seen in the crystal, offers this possibility with all four termini of a supposed dimer located on one of its sides and available for cell membrane anchoring. In order to look for evidence for at least temporary dimerization in solution, we performed chemical cross-linking using a "zero-length" condensation reagent 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC).

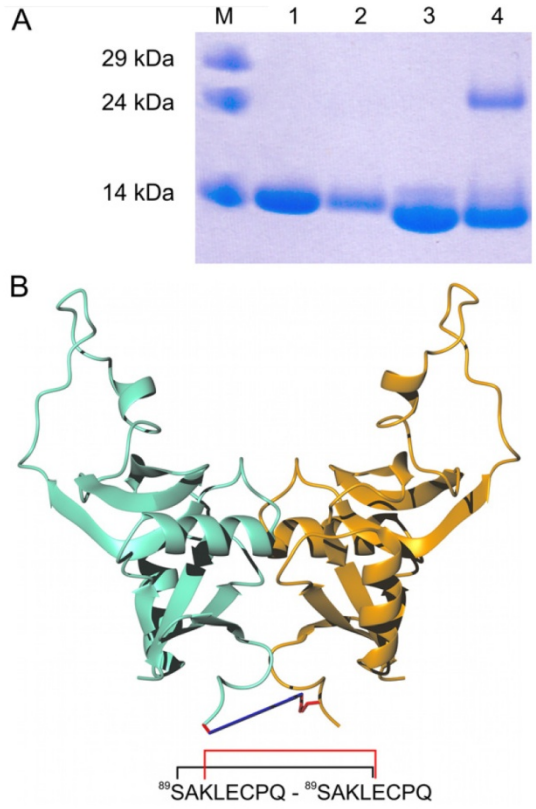


Fig.3. Formation of receptor dimers suggested by protein cross-linking. (A) Reaction mixture after chemical cross-linking analyzed by SDS–PAGE. Lane M, molecular weight markers; lane 1, lysozyme control (no EDC added); lane 2, lysozyme control with 25× molar excess of EDC; lane 3, NKR-P1Aex control (no EDC added); lane 4, mNKR-P1Aex with 25× molar excess of EDC. (B) A schematic structure of the cross-link types identified in the chemical cross-linking experiments depicted on dimer B of NKR-P1Aex.

After the cross-linking reaction, a formation of NKR-P1Aex dimers was observed, while no dimers were formed for the lysozyme control (Fig. 3A). Cross-linking occurred between the N-terminal octapeptides S^{89}AKLECPQ^{96} with carboxylate of Glu^{93} of one subunit linked through an amide bond to either the *N*-terminal amino group of Ser^{89}, or to the side chain amine of Lys^{91} of another subunit (Fig. 3B). This observation indicates a spatial proximity of the N-termini, and provides evidence for a temporary type B association.

4. Discussion

4.1. Structure of NKR-P1 receptors

The NKR-P1 family of NK cell receptors has been a subject of extensive investigations for over two decades since cloning of rat NKR-P1A (Giorda et al., 1990; Mesci et al., 2006). While extensive immunological search for the individual isoforms and their ligands has shed light onto the function of this receptor family, their production in large quantities has posed certain problems. Our previous attempts to achieve this production based on a protocol published by Kogelberg et al. (2000) have not been successful. Therefore, we have amplified the extracellular portion of mouse NKR-P1A, and produced stable and soluble protein of exceptional purity suitable for biochemical and structural investigations. The first structure of a member of the entire family allowed us to detect the characteristic features of these receptors similar to other structurally described CTLD receptors, to reveal a distinct feature of the NKR-P1 family and to confirm previous predictions (Sovova et al., 2010).

The structure of a monomer of NKR-P1Aex is very similar to its MR search model – human CD69 – except for the extended loop (residues 158–187, Fig. 2H). The compact core of NKR-P1Aex conforms to the classical CTLD fold (Zelensky and Gready, 2005). Two types of dimers observed in the crystal are both of interest: the stronger type A utilizing the swapped loops, and type B reflecting the human CD69 dimer whose presence in solution is likely (cross-linking experiments). Type A dimer has its N-terminal placed on the opposite ends of the globule (distance 67 Å). For type B this distance is shorter (28 Å), and the termini are located on one side of the dimer. In contrast, in human CD69 the *N*-termini interact with each other including several main chain–main chain hydrogen bonds resulting in short inter-chain distances, in agreement with the supposed model of dimeric stems and transmembrane domains localized close to each other on cell surface. Both A and B interactions in NKR-P1Aex crystal structure pose sufficiently strong contacts to occur *in vivo*.

Dimer B differs in the interacting residues, and mutual orientation of the monomers compared to human CD69. It places the protein termini close in space so that a membrane-proximal part of the dimer can be distinguished from the membrane-distal one with the loops. The interface interaction area and number of hydrogen bonds between the chains is similar as in mouse Ly49A and human CD94 (460 and 610 $Å^2$, respectively, Natarajan et al., 2000) and these are likely to be much larger in full-length receptor, as the N- and C-terminal parts of the receptor extracellular domain are missing from the actual NKR-P1Aex expression construct and crystal structure, respectively. Dimers of human CD69 are formed by interactions between two strands β1, and between several residues of helices α2 and α2′. In NKR-P1Aex the monomers are connected via hydrogen bonds between helix α1 and strand β6. Interactions around Arg^{207} and $Arg^{207'}$ and Gln^{110} are the most significant. Both residues are highly conserved in NKR-P1 receptors (Fig. A2), and we propose that also other members of the family can form similar dimers.

4.2. Extended loop

An extended loop pointing away from the globular CTLD was also observed in other C-type lectin like proteins important in immune system and other biochemical interactions. A well documented example in the former category is represented by the CTL domain 4 of the macrophage mannose receptor (MMR4), critical for binding of carbohydrates by this receptor (Taylor et al., 1992; Feinberg et al., 2000). However, the functional importance of this particular loop (Fig. 2E) for the biology of the macrophage mannose receptor has not been evaluated (Feinberg et al., 2000). On the other hand, the function of this structural element is well understood in snake venoms and their complexes with blood coagulation proteins. The C-type lectin heterodimeric protein from habu snake *Trimeresurus flavoviridis* dimerizes via domain swapping, which is functionally important for its interaction with blood coagulation proteins Factor IX and X (Mizuno et al., 1997, 2001). Bitiscetin, a C-type lectin snake venom inducing platelet agglutination, binds a von Willebrand factor A1 domain into the central groove of its domain-swapped dimer which is similar to type A dimer formed in the NKR-P1Aex crystal structure (Fig. 2F, Maita et al., 2003). In fact, numerous C-type lectin-like proteins of snake venom modulating the hemostatic system and forming 3D domain-swapped dimers and oligomers have been characterized up to date (reviewed in Morita (2005)). Finally, domain swapping has been described in a number of other proteins not related to C-type lectins (Liu and Eisenberg, 2002).

4.3. Ligand binding

Dimer B provides a widely open platform between the extended loops. The loops would be entirely accessible for a potential ligand. The U-shaped concave surface with a relatively flat bottom and a width of about 40 Å would provide a site mostly hydrophobic along its walls, and with two distinct patches of negative electrostatic potential near the center of its base (Fig. 2G). Ligands could bind either to the loops or become embraced by them when interacting with the site base, the latter option being a likely model for NKR-P1–Clr interactions.

As the sequence of the extended loop is largely conserved within the family of NKR-P1 receptors we suggest that other members of the family are capable of similar structural behavior. The important role of this surface loop of the CTLD fold in other proteins implies its significance in the NKR-P1 receptors where it is most likely involved in interactions with Clr ligands.

Further investigations are necessary to clarify molecular details of these "receptor"–"ligand" interactions and their relevance to those occurring at the cellular level during functionally important processes described in the immunological literature. It also remains to be established whether NKR-P1 and Clr receptors may also interact on the same cell, that is, not only in *trans* but also in *cis* mode, as was recently shown for Ly49 and MHC class I recognition system (Back et al., 2009). Even the terms "receptor" and "ligand" may be confusing in the case of the NKR-P1–Clr recognition system, given the fact that both partners possess signaling motifs and both might be expressed on each of the two cells participating in their actual immune cross-talk.

Acknowledgments

The authors wish to thank Uwe Müller for support at the beam line BL14.1 of Bessy II, Helmholtz-Zentrum Berlin, Albert-Einstein-Str. 15.

This work was supported by the Ministry of Education of the Czech Republic (MSM 21620808 and 1M0505 to K.B.), the Czech Science Foundation (303/09/0477, 305/09/H008 to K.B., 305/07/1073 to J.H. and 207/10/1040 to P.N.), the Grant Agency of the Academy of Sciences of the Czech Republic (KJB101120805 to V.K. and K.H.), and by the E.C., project SPINE2-Complexes (031220 to J.D.), and ELISA grant (226716, synchrotron access, projects 09.1.81077 and 09.2.90262).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jsb.2011.05.001.

References

Aldemir, H., Prod'homme, V., Dumaurier, M.J., Retiere, C., Poupon, G., et al., 2005. Cutting edge: lectin-like transcript 1 is a ligand for the CD161 receptor. J. Immunol. 175, 7791–7795.

Arase, N., Arase, H., Park, S.Y., Ohno, H., Ra, C., Saito, T., 1997. Association with FcRγ is essential for activation signal through NKR-P1 (CD161) in natural killer (NK) cells and NK1.1+ T cells. J. Exp. Med. 186, 1957–1963.

Arase, H., Arase, N., Saito, T., 1996. Interferon gamma production by natural killer (NK) cells and NK1.1+ T cells upon NKR-P1 cross-linking. J. Exp. Med. 183, 2391–2396.

Aust, J.G., Gays, F., Mickiewicz, K.M., Buchanan, E., Brooks, C.G., 2009. The expression and function of the NKRP1 receptor family in C57BL/6 mice. J. Immunol. 183, 106–116.

Back, J., Malchiodi, E.L., Cho, S., Scarpellino, L., Schneider, P., et al., 2009. Distinct conformations of Ly49 natural killer cell receptors mediate MHC class I recognition in *trans* and *cis*. Immunity 31, 598–608.

Baker, N.A., Sept, D., Joseph, S., Holst, M.J., McCammon, J.A., 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. Proc. Natl. Acad. Sci. USA 98, 10037–10041.

Billerbeck, E., Kang, Y.H., Walker, L., Lockstone, H., Grafmueller, S., et al., 2010. Analysis of CD161 expression on human CD8+ T cells defines a distinct functional subset with tissue-homing properties. Proc. Natl. Acad. Sci. USA 107, 3006–3011.

Boyington, J.C., Motyka, S.A., Schuck, P., Brooks, A.G., Sun, P.D., 2000. Crystal structure of an NK cell immunoglobulin-like receptor in complex with its class I MHC ligand. Nature 405, 537–543.

Carlyle, J.R., Jamieson, A.M., Gasser, S., Clingan, C.S., Arase, H., 2004. Missing self-recognition of Ocil/Clr-b by inhibitory NKR-P1 natural killer cell receptors. Proc. Natl. Acad. Sci. USA 101, 3527–3532.

Carlyle, J.R., Mesci, A., Ljutic, B., Belanger, S., Tai, L.H., et al., 2006. Molecular and genetic basis for strain-dependent NK1.1 alloreactivity of mouse NK cells. J. Immunol. 176, 7511–7524.

Cerny, J., Fiserova, A., Horvath, O., Bezouska, K., Pospisil, M., Horejsi, V., 1997. Association of human NK cell surface receptors NKR-P1 and CD94 with Src-family protein kinases. Immunogenetics 46, 231–236.

Collaborative Computational Project, Number 4, 1994. The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D 50, 760–763.

Deng, L., Cho, S., Malchiodi, E.L., Kerzic, M.C., Dam, J., Mariuzza, R.A., 2008. Molecular architecture of the major histocompatibility complex class I-binding site of Ly49 natural killer cell receptors. J. Biol. Chem. 283, 16840–16849.

Emsley, P., Cowtan, K., 2004. Coot: model building tools for molecular graphics. Acta Crystallogr. D 60, 2126–2132.

Evans, P.R., 2005. Scaling and assessment of data quality. Acta Crystallogr. D 62, 72–82.

Feinberg, H., Park-Snyder, S., Kolatkar, A.R., Heise, C.T., Taylor, M.E., Weis, W.I., 2000. Structure of a C-type carbohydrate recognition domain from the macrophage mannose receptor. J. Biol. Chem. 275, 21539–21548.

Fine, J.H., Chen, P., Mesci, A., Allan, D.S.J., Gasser, S., et al., 2010. Chemotherapy-induced genotoxic stress promotes sensitivity to natural killer cell cytotoxicity by enabling missing-self recognition. Cancer Res. 70, 7102–7113.

Giorda, R., Rudert, W.A., Vavasscori, C., Chambers, W.H., Hiserodt, J.C., Trucco, M., 1990. NKR-P1 a signal transduction molecule on natural killer cells. Science 249, 1298–1300.

Giorda, R., Trucco, M., 1991. Mouse NKR-P1. A family of genes selectively coexpressed in adherent lymphokine-activated killer cells. J. Immunol. 147, 1701–1708.

Iizuka, K., Naidenko, O.V., Plougastel, B.F., Fremont, D.H., Yokoyama, W.M., 2003. Genetically linked C-type lectin-related ligands for the NKRP1 family of natural killer cell receptors. Nat. Immunol. 4, 801–807.

Kärre, K., 2008. Natural killer cell recognition of missing self. Nat. Immunol. 9, 477–480.

Kim, S., Poursine-Laurent, J., Truscott, S.M., Lybarger, L., Song, Y.J., et al., 2005. Licensing of natural killer cells by host major histocompatibility complex class I molecules. Nature 436, 709–713.

Kleinschek, M.A., Boniface, K., Sadekova, S., Grein, J., Murphy, E.E., et al., 2009. Circulating and gut-resident human Th17 cells express CD161 and promote intestinal inflammation. J. Exp. Med. 206, 525–534.

Kogelberg, H., Lawson, A.M., Muskett, F.W., Carruthers, R.A., Feizi, T., 2000. Expression in *Escherichia coli*, folding in vitro, and characterization of the carbohydrate recognition domain of the natural killer cell receptor NKR-P1A. Protein Expr. Purif. 20, 10–20.

Kolenko, P., Skalova, T., Vanek, O., Stepankova, A., Duskova, J., et al., 2009. The high-resolution structure of the extracellular domain of human CD69 using a novel polymer. Acta Crystallogr. F 65, 1258–1260.
Krissinel, E., Henrick, K., 2004. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr. D 60, 2256–2268.
Krissinel, E., Henrick, K., 2007. Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. 372, 774–797.
Kveberg, L., Bäck, C.J., Dai, K.Z., Inngjerdingen, M., Rolstad, B., et al., 2006. The novel inhibitory NKR-P1C receptor and Ly49s3 identify two complementary, functionally distinct NK cell subsets in rats. J. Immunol. 176, 4133–4140.
Kveberg, L., Dai, K.Z., Westgaard, I.H., Daws, M.R., Fossum, S., et al., 2009. Two major groups of rat NKR-P1 receptors can be distinguished based on chromosomal localization, phylogenetic analysis and Clr ligand binding. Eur. J. Immunol. 39, 541–551.
Liu, Y., Eisenberg, D., 2002. 3D domain swapping: as domains continue to swap. Protein Sci. 11, 1285–1299.
Ljutic, B., Carlyle, J.R., Filipp, D., Nakagawa, R., Julius, M., Zuniga-Pflucker, J.C., 2005. Functional requirement for signaling through the stimulatory and inhibitory mouse NKR-P1 (CD161) NK cell receptors. J. Immunol. 174, 4789–4796.
Long, F., Vagin, A.A., Young, P., Murshudov, G.N., 2008. BALBES: a molecular-replacement pipeline. Acta Crystallogr. D 64, 125–134.
Lovell, S.C., Davis, I.W., Arendall, W.B., de Bakker, P.I.W., Word, J.M., et al., 2003. Structure validation by Cα geometry: φ, ψ and Cβ deviation. Proteins 50, 437–450.
Maita, N., Nishio, K., Nishimoto, E., Matsui, T., Shikamoto, Y., et al., 2003. Crystal structure of von Willebrand factor A1 domain complexed with snake venom, bitiscetin. J. Biol. Chem. 278, 37777–37781.
Mesci, A., Ljutic, B., Makrigiannis, A.P., Carlyle, J.R., 2006. NKR-P1 biology from prototype to missing self. Immunol. Res. 35, 13–26.
Mizuno, H., Fujimoto, Z., Atoda, H., Morita, T., 2001. Crystal structure of an anticoagulant protein in complex with the Gla domain of factor X. Proc. Natl. Acad. Sci. USA 98, 7230–7234.
Mizuno, H., Fujimoto, Z., Koizumi, M., Kano, H., Atoda, H., Morita, T., 1997. Structure of coagulation factors IX/X-binding protein, a heterodimer of C-type lectin domains. Nat. Struct. Biol. 4, 438–441.
Morita, T., 2005. Structures and functions of snake venom CLPs (C-type lectin-like proteins) with anticoagulant-, procoagulant-, and platelet-modulating activities. Toxicon 45, 1099–1114.
Murshudov, G.N., Vagin, A.A., Dodson, E.J., 1997. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D 53, 240–255.
Natarajan, K., Sawicki, M.W., Margulies, D.H., Mariuzza, R.A., 2000. Crystal structure of human CD69: a C-type lectin-like activation marker of hematopoietic cells. Biochemistry 39, 14779–14786.
Pompach, P., Man, P., Kavan, D., Hofbauerova, K., Kumar, V., et al., 2009. Modified electrophoretic and digestion conditions allow a simplified mass spectrometric evaluation of disulfide bonds. J. Mass Spectrom. 44, 1565–1570.
Rosen, D.B., Bettadapura, J., Alsharifi, M., Mathew, P.A., Warren, H.S., et al., 2005. Cutting edge: lectin-like transcript-1 is a ligand for the inhibitory human NKR-P1A receptor. J. Immunol. 175, 7796–7799.
Rosen, D.B., Cao, W., Avery, D.T., Tangye, S.G., Liu, Y.J., et al., 2008. Functional consequences of interactions between human NKR-P1A and its ligand LLT1 expressed on activated dendritic cells and B cells. J. Immunol. 180, 6508–6517.
Roth, P., Mittelbronn, M., Wick, W., Meyermann, R., Tatagiba, M., Weller, M., 2007. Malignant glioma cells counteract antitumor immune responses through expression of lectin-like transcript-1. Cancer Res. 67, 3540–3544.
Ryan, J.C., Niemi, E.C., Goldfien, R.D., Hiserodt, J.C., Seaman, W.E., 1991. NKR-P1, an activating molecule on rat natural killer cells, stimulated phosphoinositide turnover and a rise in intracellular calcium. J. Immunol. 147, 3244–3250.
Ryan, J.C., Niemi, E.C., Nakamura, M., Seaman, W.E., 1995. NKR-P1A is a target specific receptor that activates natural killer cell cytotoxicity. J. Exp. Med. 181, 1915–1991.
Ryan, J.C., Seaman, W.E., 1997. Divergent functions of lectin-like receptors of NK cells. Immunol. Rev. 155, 79–89.
Schuck, P., 2000. Size distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modelling. Biophys. J. 78, 1606–1619.
Schuck, P., 2003. On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation. Anal. Biochem. 320, 104–124.
Skalova, T., Duskova, J., Hasek, J., Kolenko, P., Stepankova, A., Dohnalek, J., 2010. Alternative polymer precipitants for protein crystallization. J. Appl. Cryst. 43, 737–742.
Sovova, Z., Kopecky Jr., V., Pazderka, T., Hofbauerova, K., Rozbesky, D., et al., 2010. Structural analysis of natural killer cell receptor protein 1 (NKR-P1) extracellular domains suggests a conserved long loop region involved in ligand specificity. J. Mol. Model. doi:10.1007/s00894-010-0837-y.
Taherzadech, Z., VanBavel, E., de Vos, J., Matlung, H.L., van Vontrans, G., et al., 2010. Strain-dependent susceptibility for hypertension in mice resides in the natural killer gene complex. Am. J. Physiol. Heart Circ. Physiol. 298, H1273–H1282.
Taylor, M.E., Bezouška, K., Drickamer, K., 1992. Contribution to ligand binding by multiple carbohydrate-recognition domain in the macrophage mannose receptor. J. Biol. Chem. 267, 1719–1726.
Tian, W., Feng, B., Liou, H.C., 2005a. Silencing OCILRP2 leads to intrinsic defects in T cells in response to antigenic stimulation. Cell. Immunol. 235, 72–84.
Tian, W., Nunez, R., Cheng, S., Ding, Y., Tumang, J., 2005b. C-type lectin OCILRP2/Clr-g and its ligand NKRP1f costimulate T cell proliferation and IL-2 production. Cell. Immunol. 234, 39–53.
Vaguine, A.A., Richelle, J., Wodak, S.J., 1999. SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with atomic model. Acta Crystallogr. D 55, 191–205.
Valiante, N.M., Lienert, K., Shilling, H.G., Smits, B.J., Parham, P., 1997. Killer cell receptors: keeping pace with MHC class I evolution. Immunol. Rev. 155, 155–164.
Vanek, O., Nalezkova, M., Kavan, D., Borovickova, I., Pompach, P., et al., 2008. Soluble recombinant CD69 receptors optimized to have an exceptional physical and chemical stability display prolonged circulation and remain intact in the blood of mice. FEBS J. 275, 5589–5606.
Vivier, E., Tomasello, E., Baratin, M., Waltzer, T., Ugolini, S., 2008. Functions of natural killer cells. Nat. Immunol. 9, 503–510.
Voigt, S., Mesci, A., Ettinger, J., Fine, J.H., Chen, P., et al., 2007. Cytomegalovirus evasion of innate immunity by subversion of the NKR-P1B:Clr-b missing-self axis. Immunity 26, 617–627.
Yang, T., Flint, M.S., Webb, K.M., Chambers, W.H., 2006. CD161B:ClrB interactions mediate activation of enhanced lysis of tumor target cells following NK cell:DC co-culture. Immunol. Res. 36, 43–50.
Zelensky, A.N., Gready, J.E., 2005. The C-type lectin-like domain superfamily. FEBS J. 272, 6179–6217.

Table A1
Assignment of the Raman vibrational bands distinguishable in the spectra of NKR--P1Aex. The following symbols are used: ν = stretching vibrations, δ = bending vibrations, *sh* = shoulder, *w* = weak.

Wavenumber (cm^{-1})	Assignment	Reference[a]
509	νS–S of Cys in GGG	1
536	νS–S of Cys in TGT	1
570	Trp	1
621	Phe	2–4
643	Tyr	2–4
656 *sh*	νC–S of Cys in P_H and Met in P_C	1
710 *w*	νC–S of Cys in P_C and P_N, Met in P_C	1
760	Trp W18	2–5
829	Tyr doublet, Phe	1–5
853	Tyr doublet, Ile	1–5
877	Trp W17, Ile, Val	2–5
901 *sh*	Lys, Ala	2, 4
940 *sh*	νC–C (α-helix), Lys, Val, Leu	2, 4
961	νC–C (irregular), Lys, Leu	2, 4
986 *sh*	Ile	2, 4
1003	Phe	1–4
1010 *sh*	Trp shoulder	1–4
1031	Phe, Tyr	1–4
1080	Lys, Phe	2–4
1128	Ile, Val, Leu, Trp	2–4
1155 *sh*	Ile, Val	2–4
1177	Tyr, Phe	2–4
1208	Phe, Tyr, Trp	2–4
1235	amide III (α-helix)	1–5
1256	amide III (irregular/β-turn), Trp	1–5
1311	amide III (α-helix/β-turn), Lys, Val, Ile	1–5
1340	Trp doublet	1–5
1359 *sh*	Trp doublet, Val, Phe	1–5
1399	His, νCO_2^- of Asp, Glu	4
1444	δCH_2, δCH_3, Lys, Ile, Leu	2–4
1459	δCH_2, Ala, Ile, Val, Leu, Trp, Tyr	2–4
1495 *w*	Trp, Phe	2–4
1551	Trp W3	4
1578	Trp, Phe, Tyr	2–4
1586 *sh*	Trp, Phe, Tyr	2–4
1607	Phe, Tyr Y8b	2–5
1619	Tyr Y8a, Trp	2–5
1669	amide I	1–5

[a] The used references are: 1) Tensmeyer and Kauffman, 1996; 2) Overman and Thomas Jr., 1999; 3) Overman and Thomas Jr., 1995; 4) Tuma, 2005; 5) Miura and Thomas Jr., 1995.

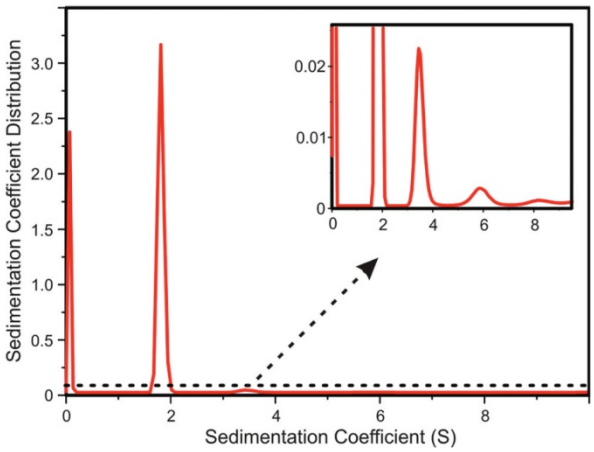


Fig.A1. Sedimentation velocity analysis of NKR-P1Aex. The majority of NKR-P1Aex behaves as a stable monomer corresponding to the most intensive peak in the c(s) profile but a more detailed view (insert) reveals the presence of several oligomeric species which were also detected in the chemical cross-linking experiment.

Table A2
Estimates of the secondary structure content of mNKR-P1Aex in aqueous solution and in the crystal from Raman spectra. The secondary structure content is given as percentage values, based on the pattern recognition least-squares method (LSA; Williams, 1986) and on the reference intensity profiles method (4-RIP; Berjot et al., 1987) that analyze amide I region. Standard deviations are calculated as standard deviations of the used reference sets. The numbers in parentheses correspond to number of residues of a given secondary structure.

	SOLUTION		CRYSTAL	
Structure	LSA	4-RIP	LSA	4-RIP
α-helix	30 ± 5 (41)	26 ± 3 (36)	29 ± 5 (41)	26 ± 3 (36)
α-ordered	22 ± 4 (30)	14 ± 3 (19)	20 ± 4 (29)	15 ± 3 (21)
α-disordered	8 ± 4 (11)	12 ± 3 (17)	9 ± 4 (12)	11 ± 3 (15)
β-sheet	42 ± 4 (58)	47 ± 3 (65)	42 ± 4 (58)	44 ± 3 (62)
β-turn	18 ± 2 (26)	—	19 ± 2 (26)	—
Unordered	10 ± 2 (14)	27 ± 3 (38)	10 ± 2 (14)	30 ± 3 (41)

Fig.A2. Multiple sequence alignment of the NKR-P1 receptor family. Only the conserved C-type lectin-like domain sequences are shown. Conservation of the NKR-P1Aex extended loop sequence is highlighted in green. Blue numbering identifies Cys residues involved in disulfide bridges by the same number. Sequence of hCD69, the MR search model and the closest known structure, and of mClrb was added manually. MusA mouse NKR-P1A, musB mouse NKR-P1B, musC mouse NKR-P1C, NK1.1 mouse NK1.1 alloantigen, musD mouse NKR-P1D, ratA rat NKR-P1A, ratB rat NKR-P1B, musF mouse NKR-P1F, musG mouse NKR-P1G, ratF rat NKR-P1F, equB horse NKR-P1B, humA human NKR-P1A, CD69 human CD69, CLRB mouse Clrb. Alignment of the NKR-P1 proteins was performed by program BLAST (Altschul et al., 1990).

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Berjot, M., Marx, J., Alix, A.J.P., 1987. Determination of the secondary structure of proteins from the Raman amide I band: the reference intensity profiles method. J. Raman Spectrosc. 18, 289–300.

Miura, T., Thomas Jr., G.J., 1995. Raman spectroscopy of proteins and their assemblies, in: Biswas, B.B., Roy, S. (Eds.), Subcellular Biochemistry, vol. 24, Proteins: Structure, Function, and Engineering. Plenum Press, New York, pp. 55–99.

Overman, S.A., Thomas Jr., G.J., 1995. Raman spectroscopy of the filamentous virus *Ff* (*fd*, *fl*, *Ml3*): structural interpretation for coat protein aromatics. Biochemistry 34, 5440–5451.

Overman, S.A., Thomas Jr., G.J., 1999. Raman markers of nonaromatic side chains in an α-helix assembly: Ala, Asp, Glu, Gly, Ile, Leu, Lys, Ser, and Val residues of phage *fd* subunits. Biochemistry 38, 4018–4027.

Tensmeyer, L.G., Kauffman II, E.W., 1996. Protein structure as revealed by nonresonance Raman spectroscopy, in: Havel, H.A. (Ed.), Spectroscopic Methods for Determining Protein Structure in Solution. VCH Publishers, New York, pp. 69–95.

Tuma, R., 2005. Raman spectroscopy of proteins: from peptides to large assemblies. J. Raman Spectrosc. 36, 307–319.

Williams, R.W., 1986. The secondary structure analysis using Raman amide I and amide III spectra. Methods Enzymol. 130, 311–331.

PAPER IV

Kolenko, P., Rozbesky, D., Vanek, O., Bezouska, K., Hasek, J., Dohnalek, J

Structure of the H107R variant of the extracellular domain of mouse NKR-P1A at 2.3 A resolution.

Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. **67**, 1519-23 (2011).

My contribution to the publication: *research performing (cloning, protein expression, refolding in vitro, purification), data collection, data analysis and interpretation, manuscript writing*

Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Structure of the H107R variant of the extracellular domain of mouse NKR-P1A at 2.3 Å resolution

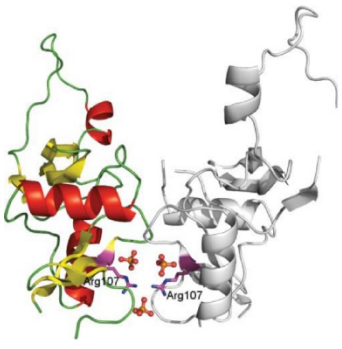
Petr Kolenko,[a]* Daniel Rozbeský,[b,c] Ondřej Vaněk,[b,c] Karel Bezouška,[b,c] Jindřich Hašek[a] and Jan Dohnálek[a,d]

[a]Institute of Macromolecular Chemistry AS CR, v.v.i., Heyrovského nám. 2/1888, 162 06 Prague 6, Czech Republic, [b]Department of Biochemistry, Faculty of Science, Charles University, Hlavova 8, 12840 Prague 2, Czech Republic, [c]Institute of Microbiology AS CR, v.v.i., Vídeňská 1083, 142 20 Prague 4, Czech Republic, and [d]Institute of Physics AS CR, v.v.i., Na Slovance 2, 182 21 Prague 6, Czech Republic

Correspondence e-mail: kolenko@imc.cas.cz

Received 23 September 2011
Accepted 2 November 2011

PDB Reference: H107R mNKR-P1A, 3t3a.



The structure of the H107R variant of the extracellular domain of the mouse natural killer cell receptor NKR-P1A has been determined by X-ray diffraction at 2.3 Å resolution from a merohedrally twinned crystal. Unlike the structure of the wild-type receptor in space group $I4_122$ with a single chain per asymmetric unit, the crystals of the variant belonged to space group $I4_1$ with a dimer in the asymmetric unit. Different degrees of merohedral twinning were detected in five data sets collected from different crystals. The mutation does not have a significant impact on the overall structure, but led to the binding of an additional phosphate ion at the interface of the molecules.

1. Introduction

Natural killer (NK) cells play an important role in the innate immune response against viruses, parasites, intracellular bacteria and tumour cells (Makrigiannis *et al.*, 2001; Bezouška *et al.*, 1994). NK cells eliminate their targets using two distinct mechanisms: cytokine secretion and cell-mediated cytotoxicity (Mesci *et al.*, 2006). The activity towards the target cells is regulated by a large number of receptors on the cell surface (Vivier *et al.*, 2008).

The activating and inhibitory receptors of NK cells fall into two distinct structural classes: those belonging to the immunoglobulin superfamily and those of the C-type lectin-like (CTL) type. The receptors of the NKR-P1 family belong to the CTL type (Kolenko *et al.*, 2011). When an NK cell encounters a target cell, the NKR-P1 receptors are expected to engage in interactions with their protein partners expressed on the surface of the target. Although the protein ligands have been identified as cell-surface CTL proteins, the details of the interactions as well as the exact route of NKR-P1 signalling are not yet fully understood. Recent studies have supported the participation of NKR-P1 receptors in protein–protein interactions with C-type lectin-related (Clr) molecules, *e.g.* LLT1 is a ligand of human NKR-P1 (Aldemir *et al.*, 2005; Rosen *et al.*, 2005) and Clrb/Ocil binds mouse NKR-P1B and NKR-D (Iizuka *et al.*, 2003; Carlyle *et al.*, 2004; Kveberg *et al.*, 2009). The function of the remaining receptors is predicted from the amino-acid composition of the intracellular part, *i.e.* whether it contains the immunoreceptor tyrosin-based inhibitory motif (ITIM) or activation motif (ITAM) (Mesci *et al.*, 2006).

The recently determined crystal structure of the extracellular domain of the mouse NKR-P1A receptor (mNKR-P1A) was the first structure of a representative of the NKR-P1 family. mNKR-P1A is formed by a compact CTL core and an extended loop that participates in domain swapping. A potential role of the swapped loop in natural ligand binding has been suggested by *in silico* studies (Sovová *et al.*, 2011). Even if a protein ligand (or ligands) for mNKR-P1A were suggested, firm proof of the interactions on the molecular level and of complex formation and consequent signalling-cascade events is still lacking.

In a series of cloning and protein-expression experiments, a mutant H107R was unintentionally generated which could be purified to sufficient purity and homogeneity and crystallized. In the process of structure solution the mutation of residue 107 was observed and was confirmed by sequencing. The mutation causes interesting behaviour

Table 1
Data-collection statistics.

Values in parentheses are for the outer resolution shell.

X-ray source†	BESSY II	IP	IP	IP	IP
Wavelength (Å)	0.918	1.542	1.542	1.542	1.542
Resolution (Å)	50.0–2.3 (2.42–2.30)	40.8–2.6 (2.67–2.60)	30.0–3.0 (3.16–3.00)	40.5–3.1 (3.27–3.10)	40.0–3.5 (3.62–3.50)
Space group	$I4_1$	$I4_1$	$I4_1$	$I4_1$	$I4_1$
Unit-cell parameters (Å)					
$a = b$	64.7	64.3	64.0	64.4	64.1
c	156.1	158.1	157.0	156.1	157.5
No. of observations	107152 (15469)	37107 (2658)	20972 (2880)	14941 (2168)	12419 (1223)
No. of unique reflections	14248 (2064)	9801 (719)	6301 (908)	5417 (804)	3933 (387)
Completeness (%)	100.0 (100.0)	99.4 (98.8)	99.5 (99.6)	99.5 (95.0)	97.8 (99.2)
Multiplicity	7.5 (7.5)	3.8 (3.7)	3.3 (3.2)	2.8 (2.7)	3.2 (3.2)
$\langle I/\sigma(I)\rangle$	16.0 (3.4)	9.4 (2.5)	10.0 (2.5)	7.6 (2.4)	9.2 (3.3)
Wilson B factor (Å^2)	49	76	71	80	46
R_{merge}	0.072 (0.546)	0.049 (0.696)	0.107 (0.460)	0.100 (0.406)	0.120 (0.321)
α_{twin}	0.14	0.09	0.26	0.19	0.21

† BESSY II, beamline 14.1 at the BESSY II synchrotron-radiation source, Helmholtz-Zentrum Berlin; IP, Agilent Technologies Gemini Enhanced Ultra sealed tube at Institute of Physics AS CR, v.v.i., Prague.

Table 2
Refinement statistics for the synchrotron data set.

Resolution (Å)	50.0–2.3
No. of residues in the asymmetric unit	241
Localized ions	3 PO_4^{3-}
No. of water molecules	56
No. of localized non-H atoms	2016
Mean ADP for protein atoms (Å^2)	45
Mean ADP for solvent (Å^2)	48
Mean ADP for all atoms (Å^2)	45
R.m.s.d. bonds (Å)	0.017
R.m.s.d. angles (°)	1.620
Ramachandran plot	
No. of residues in favoured region	224
No. of residues in allowed region	9
No. of residues in outlier region	0
R_{work}	0.188
R_{free}	0.216
No. of reflections for R_{free} calculation	708 (5%)
R_{all}	0.191

of the receptor dimers, affects the symmetry of the dimer interface and gives rise to twinning.

2. Methods

2.1. Protein preparation and crystallization

The plasmid for the expression of mNKR-P1A was constructed as described previously (Kolenko *et al.*, 2011). DNA-sequence analysis of the constructed plasmid revealed a novel point mutation in the gene for NKR-P1A leading to the substitution of His107 by Arg. Recombinant H107R variant was expressed and purified as described previously. Briefly, the plasmid was transformed into *Escherichia coli* BL21 (DE3) Gold (Stratagene). The protein was produced as inclusion bodies and refolded *in vitro*. The refolded protein was purified using ion-exchange chromatography on Q Sepharose FF (1.0 × 6.0 cm) and Superdex 75 HR 10/300 GL columns (GE Healthcare). The monodispersity of the monomeric protein solution was further assessed by dynamic light-scattering and SDS–PAGE analysis.

Prior to crystallization, the protein was concentrated to 10 mg ml^{-1} in a buffer solution consisting of 50 m*M* NaCl, 1 m*M* NaN_3, 10 m*M* HEPES pH 7.5. The search for a novel crystallization condition was also performed using the PolyA and PolyB screens (Skálová *et al.*, 2010). However, the only well diffracting crystals grew from 0.3 *M* ammonium phosphate, which was the original crystallization condition for wild-type mNKR-P1A. Tetragonal bipyramids appeared after 4 d at a temperature of 291 K. Prior to diffraction experiments, the crystals were flash-cooled in liquid nitrogen with 20% ethylene glycol as a cryoprotectant.

2.2. Data collection and structure determination

Four diffraction data sets were collected from different crystals on a Gemini Enhanced Ultra diffractometer at copper wavelength using an Atlas CCD detector (Agilent Technologies) at the Institute of Physics AS CR, v.v.i., Prague. The data set with the highest resolution was collected on beamline 14.1 of BESSY II at the Helmholtz-Zentrum Berlin. The synchrotron data set was collected at 100 K; all other data sets were collected at 120 K. Integration of the diffraction images was performed using the *XDS* program package (Kabsch, 2010) and scaling was performed using *SCALA* from the *CCP*4 program package (Winn *et al.*, 2011). The crystals belonged to space group $I4_1$. Analysis of the diffraction data revealed variable merohedral twinning of all crystals ($\alpha_{twin} \simeq 0.09$–$0.26$, twin law $k, h, -l$). The data-processing statistics are given in Table 1.

The structure was solved by molecular replacement with the program *MOLREP* (Vagin & Teplyakov, 2010) using the structure of wild-type mNKR-P1A as a search model (PDB entry 3m9z; Kolenko *et al.*, 2011). Structure refinement against experimental intensities with automated twin detection was carried out with *REFMAC*5 (Murshudov *et al.*, 2011) using a built-in twin-refinement protocol. Manual corrections of the model were performed with the program *Coot* (Emsley & Cowtan, 2004).

Structure validation was carried out with *MolProbity* (Chen *et al.*, 2010). All residues of the model were found in the allowed regions of the Ramachandran plot. The statistics of structure refinement are given in Table 2. The coordinates and structure factors have been deposited in the PDB (Berman *et al.*, 2000) with access code 3t3a.

3. Results and discussion

3.1. Overall structure

We produced and crystallized the H107R variant of the extracellular domain of the mouse NK-cell activation receptor NKR-P1A. The crystals belonged to space group $I4_1$ with merohedral twinning. In a search for a sample that would provide untwinned data, five data sets were collected. However, twinning with a variable twin factor was repeatedly observed. The structure was determined by molecular replacement using the extracellular domain of wild-type mNKR-P1A as a search model. The structure contained two protein chains consisting of residues Glu93–Leu214 with three phosphate ions attached

at the interface of the molecules (see Fig. 1). The quality of the observed electron density did not allow the localization of all of the atoms of residues Tyr149, Pro161 and Asp162 in chain *B*. Two phosphate ions are bound close to the interface between the monomers in the asymmetric unit (similar to the structure of the wild type) and a third phosphate ion with partial occupancy was localized between Arg107 residues belonging to opposite chains. The latter participates in interactions with different side chains from the two monomers (see Fig. 1*b*) and represents some asymmetry in this region.

The two protein chains are highly similar in structure (r.m.s.d. of 1.7 Å for C^{α} atoms, secondary-structure matching superposition algorithm) and each is formed by a compact core region similar to most C-type lectins (two main α-helices and two antiparallel β-sheets stabilized by three disulfide bridges) and a loop consisting of 30 residues (Tyr158–Asp187) which exhibits domain swapping with a symmetry-related molecule.

3.2. Interactions within the crystal lattice

Surprisingly, a more extensive interaction was found between two symmetry-related molecules within the crystal lattice than between the monomers in the asymmetric unit. This interaction corresponds to the domain-swapping effect of the extended loops Tyr158–Asp187, with a total interface area of about 1700 $Å^2$ involving hydrogen bonds and hydrophobic interactions. The potential participation of the loop in interactions with Clr molecules has been suggested, but has not yet been structurally confirmed (Kolenko *et al.*, 2011). The mutated residue Arg107 is not localized in this region, so the mutation at this residue could not affect the formation of the domain-swapped dimers observed in the wild-type structure.

The monomers in the asymmetric unit are related by a noncrystallographic twofold axis and form a noncovalent dimer with an interface area of 730 Å (Krissinel & Henrick, 2007). The difference in the size of the asymmetric unit of the wild-type structure and the mutant structure is exactly reflected by the presence of the additional twofold symmetry (monomer and dimer in the asymmetric unit; space groups $I4_122$ and $I4_1$, respectively). The twinning observed in the case of the H107R variant also corresponds to the 'missing' twofold symmetry with respect to the wild-type crystals. The arrangement of the molecules in the asymmetric unit of the H107R variant has previously been confirmed to be a potential biological unit by cross-linking experiments (Kolenko *et al.*, 2011). The mutated residue Arg107 is located at this interface and together with the partially occupied phosphate ion contributes to stabilization of the interface (see Fig. 1*b*). The His107 residues in the wild-type structure are too distant from each other to participate in any interaction.

3.3. H107R mutant *versus* wild type

The overall structure of the monomers does not differ significantly from the wild-type structure (Fig. 2). The r.m.s.d. between C^{α} atoms of the wild type and chains *A* and *B* of the mutant are 0.6 and 1.8 Å, respectively. The more similar chain *A* of the mutant contains only one region, consisting of seven residues (174–180), in which the differences between C^{α} atoms exceed 1 Å. The most significant structural differences in the mutated chain *B* are observed in the region 148–150 (in which the differences exceed 3 Å) and in a large C-terminal region containing residues 163–205, in which the differences between equivalent C^{α} atoms exceed 2 Å for the majority of superimposed pairs. However, these parts of the protein are not involved in the formation of the interface containing residue 107. Nevertheless, the mutation and binding of phosphate ion at the interface probably plays a role in the twinning of the crystals of the H107R mutant.

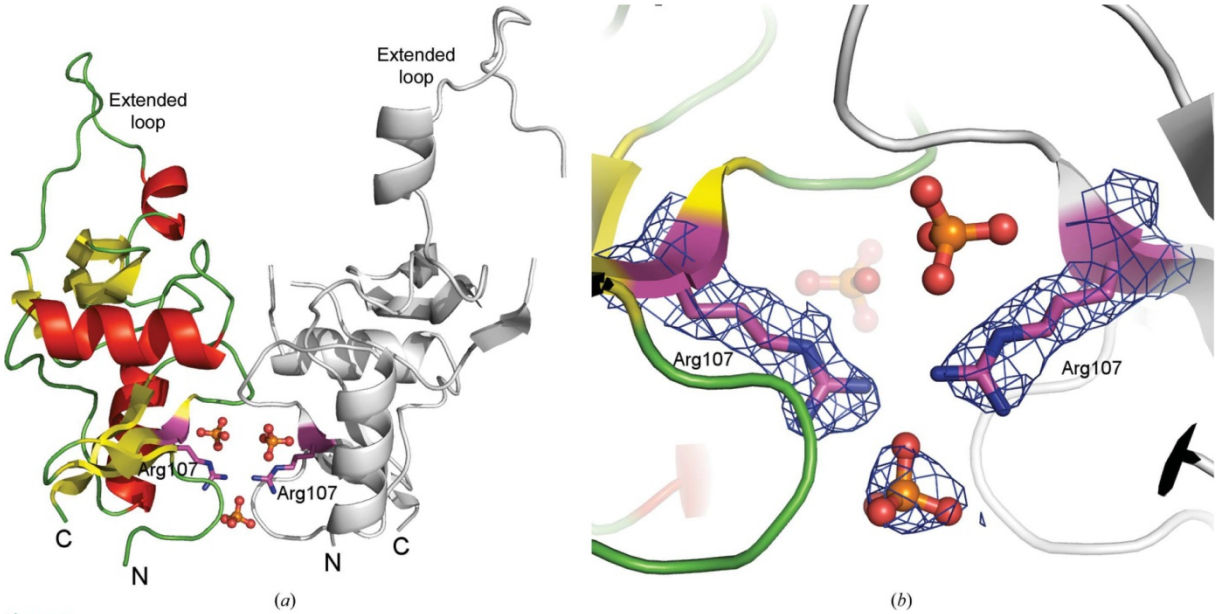


Figure 1
Overall structure of the H107R variant of mouse NKR-P1A and a detailed view of the interface between monomers. (*a*) Chain *A* is represented by coloured secondary-structure elements, chain *B* is shown in grey, phosphate ions are shown in ball-and-stick representation and side chains of Arg107 are represented by sticks. (*b*) A detailed view of the interface between the monomers and the binding of the additional phosphate ion. The $2F_o - F_c$ map around the phosphate and Arg107 is contoured at the 1σ level (blue).

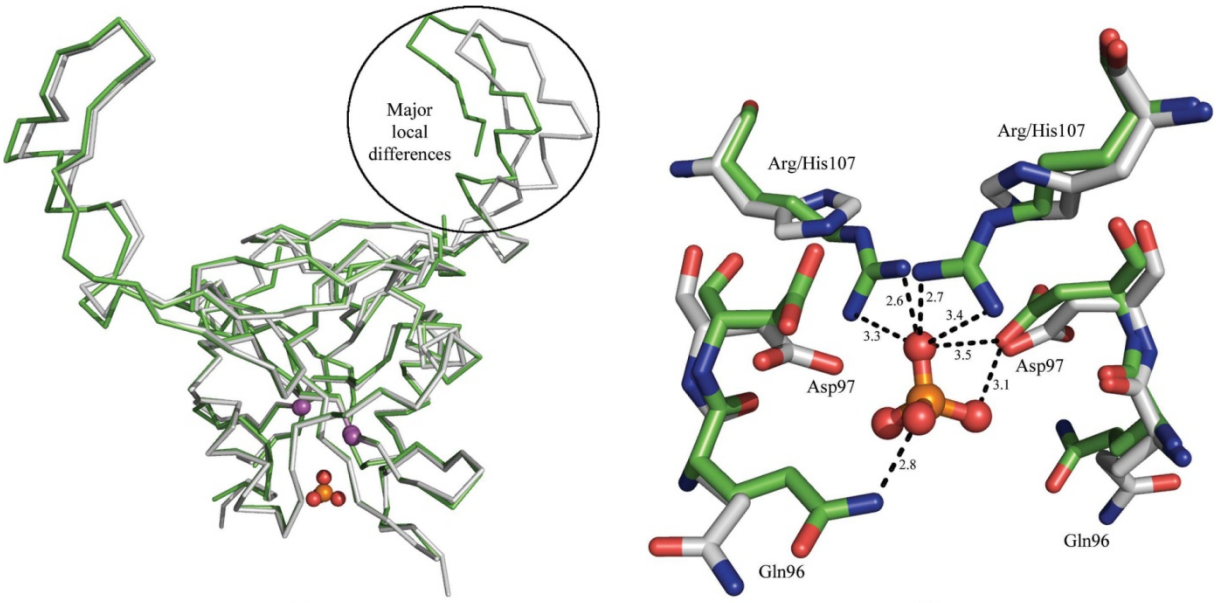


Figure 2
Structural alignment of dimers of H107R-variant (green) and wild-type (grey) mNKR-P1A. (*a*) Chains are represented by C^{α} traces, the C^{α} atoms of Arg/His107 are shown as magenta spheres and the additional phosphate ion is shown in ball-and-stick representation. The region of major local differences, the extended loop (chain *B* in the H107R variant), is highlighted by an oval. (*b*) A detailed view of the mutation-induced asymmetric binding of the phosphate ion compared with the wild type. Amino acids are represented by sticks and the phosphate ion is shown in ball-and-stick representation; the hydrogen-bonding network with distances in Å is only displayed for the discussed phosphate ion.

We analyzed the interface containing residues 107 using the *PISA* server (Krissinel & Henrick, 2007). In the cystal structure of the wild-type receptor, a buried surface area of about 650 Å^2 containing 11 hydrogen bonds was observed. The interface of the H107R mutant variant contains 13 hydrogen bonds, with a buried surface of about 730 Å^2. Contrary to our expectations, this analysis suggests that the H107R mutation caused a strengthening of the interface of interest.

Although many trials to improve the quality of the crystals were performed, the highest resolution achieved for the H107R variant (2.3 Å) is not sufficient for interpretation of the fine structural differences that lead to merohedral twinning in this case. Our interpretation of the observed phenomenon is as follows. The asymmetric binding of the phosphate ions causes a loss of the perfect symmetry in dimers of mNKR-P1A, which are then organized in two possible orientations in the crystal owing to a very high structure similarity between the monomers. The ratio between the orientations is random (varying twin factor in different crystals). The structure of the mutant in the absence of phosphate would be of particular interest. Unfortunately, despite an extensive search such conditions were not found. The present structure shows how small structural differences may lead to defects in crystal growth that affect structure analysis.

The occurrence of nonsymmetrical dimers in the asymmetric unit of the crystal unfortunately does not imply their presence in solution or their natural role in the NK cells, although to some extent it increases support for the natural occurrence of such dimers when previously confirmed by alternative methods (Kolenko *et al.*, 2011).

The authors wish to acknowledge the support of the EC Integrated project SPINE2-Complexes (LSHG-CT-2006-031220), ELISA grant No. 226716 (synchrotron-access funding, projects 09.1.81077 and 09.2.90262), the Academy of Sciences (Praemium Academiae), the Czech Science Foundation (project Nos. 305/07/1073 and P302/11/0855), the Ministry of Education of the Czech Republic (MSM 21620808 and 1M0505) and Charles University Prague (263209/2011 and 403211/2011).

References

Aldemir, H., Prod'homme, V., Dumaurier, M.-J., Retiere, C., Poupon, G., Cazareth, J., Bihl, F. & Braud, V. M. (2005). *J. Immunol.* **175**, 7791–7795.
Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.
Bezouška, K., Yuen, C.-T., O'Brien, J., Childs, R. A., Chai, W., Lawson, A. M., Drbal, K., Fišerová, A., Pospíšil, M. & Feizi, T. (1994). *Nature (London)*, **372**, 150–157.
Carlyle, J. R., Jamieson, A. M., Gasser, S., Clingan, C. S., Arase, H. & Raulet, D. H. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 3527–3532.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Iizuka, K., Naidenko, O. V., Plougastel, B. F., Fremont, D. H. & Yokoyama, W. M. (2003). *Nature Immunol.* **4**, 801–807.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Kolenko, P., Rozbeský, D., Vaněk, O., Kopecký, V., Hofbauerová, K., Novák, P., Pompach, P., Hašek, J., Skálová, T., Bezouška, K. & Dohnálek, J. (2011). *J. Struct. Biol.* **175**, 434–441.
Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
Kveberg, L., Dai, K. Z., Westgaard, I. H., Daws, M. R., Fossum, S., Naper, C. & Vaage, J. T. (2009). *Eur. J. Immunol.* **39**, 541–551.
Makrigiannis, A. P., Pau, A. T., Saleh, A., Winkler-Pickett, R., Ortaldo, J. R. & Anderson, S. K. (2001). *J. Immunol.* **166**, 5034–5043.
Mesci, A., Ljutic, B., Makrigiannis, A. P. & Carlyle, J. R. (2006). *Immunol. Res.* **35**, 13–26.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Rosen, D. B., Bettadapura, J., Alsharifi, M., Mathew, P. A., Warren, H. S. & Lanier, L. L. (2005). *J. Immunol.* **176**, 7796–7799.
Skálová, T., Dušková, J., Hašek, J., Kolenko, P., Štěpánková, A. & Dohnálek, J. (2010). *J. Appl. Cryst.* **43**, 737–742.
Sovová, Z., Kopecký, V., Pazderka, T., Hofbauerová, K., Rozbeský, D., Vaněk, O., Bezouška, K. & Ettrich, R. (2011). *J. Mol. Model.* **17**, 1353–1370.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.
Vivier, E., Tomasello, E., Baratin, M., Walzer, T. & Ugolini, S. (2008). *Nature Immunol.* **9**, 503–510.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.

PAPER V

Rozbesky, D., Man, P., Kavan, D., Chmelik, J., Cerny, J., Bezouska, K., Novak, P.

Chemical Cross-Linking and H/D Exchange for Fast Refinement of Protein Crystal Structure.

Anal. Chem. **84**, 867-870 (2012)

My contribution to the publication: *research design, research performing (protein expression, refolding in vitro, purification, chemical cross-linking, H/D exchange), data collection, data analysis and interpretation, manuscript writing*

analytical chemistry

Letter

pubs.acs.org/ac

Chemical Cross-Linking and H/D Exchange for Fast Refinement of Protein Crystal Structure

Daniel Rozbesky,[†,‡] Petr Man,[†,‡] Daniel Kavan,[†,‡] Josef Chmelik,[‡] Jiri Cerny,[§] Karel Bezouska,[†,‡] and Petr Novak*[,†,‡]

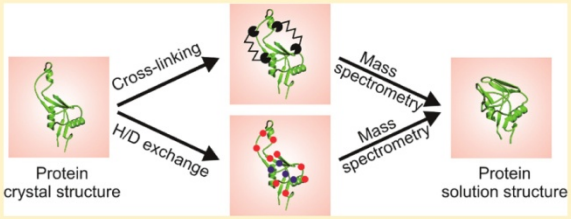
[†]Department of Biochemistry, Faculty of Science, Charles University in Prague, Prague, Czech Republic

[‡]Institute of Microbiology, Academy of Sciences of the Czech Republic, Prague, Czech Republic

[§]Institute of Biotechnology, Academy of Sciences of the Czech Republic, Prague, Czech Republic

Supporting Information

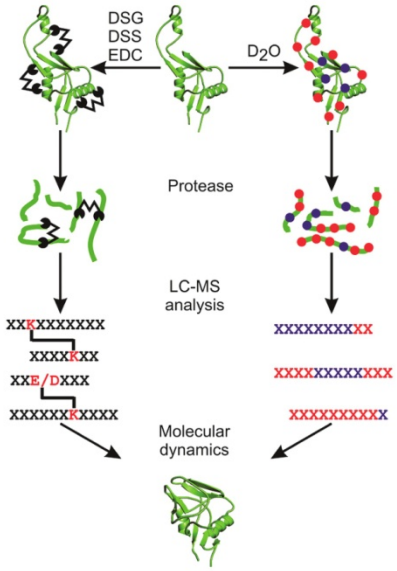
ABSTRACT: A combination of chemical cross-linking and hydrogen−deuterium exchange coupled to high resolution mass spectrometry was used to describe structural differences of NKR-P1A receptor. The loop region extended from the compact core in the crystal structure was found to be closely attached to the protein core in solution. Our approach has potential to refine protein structures in solution within a few days and has very low sample consumption.



Determining a protein structure is a prerequisite for fully understanding its function. The three-dimensional structure of proteins has been traditionally solved by X-ray crystallography or NMR spectroscopy. These techniques provide high resolution atomic data on proteins and their complexes but have some notable limitations. X-ray crystallography requires diffraction-quality protein crystals, and the structure determined by this method may be influenced by the crystal lattice. Furthermore, the crystal structure provides only a static picture of a protein under nonphysiological conditions. In contrast to X-ray crystallography, NMR spectroscopy affords dynamic structural information in solution, as well as thermodynamic and kinetic information about a protein structure. On the other hand, this method is currently restricted to proteins of molecular weight not exceeding 50 kDa. Moreover, both X-ray crystallography and NMR spectroscopy require large amounts of pure protein and are laborious.

One of the challenges of current structural proteomics is to implement high-throughput and robust techniques to probe protein structure. Mass spectrometry has become a key player in high-throughput proteomics.[1] In a perspective of *Nature Methods* in 2008, Albert Heck[2] presented an idea that future native mass spectrometry studies should focus more on the structure, topology, and dynamics of interesting proteins and protein complexes. Here, we introduce a combination of chemical cross-linking, hydrogen/deuterium (H/D) exchange and high resolution mass spectrometry to elucidate structural details of the receptor NKR-P1A (CD161a) (Scheme 1). This protein is an important activating receptor expressed at the surface of natural killer (NK) cells that act against virally infected and tumor cells.[3] The recently determined crystal structure of the extracellular domain of NKR-P1A[4] raises questions about the unique conformation of the loop region containing residues T159-D187.

Scheme 1. General Strategies for Protein Structure Characterization by Chemical Cross-Linking (Left) and H/D Exchange (Right) Coupled to Mass Spectrometry



In a homology model[5] as well as in structures of homologous proteins, this evolutionarily conserved loop is in close proximity to the compact core. Surprisingly, this loop was found to be

Received: October 25, 2011
Accepted: December 15, 2011
Published: December 15, 2011

extended from the compact core in the crystal structure of NKR-P1A (Figure 1a) and bound to a symmetry-related

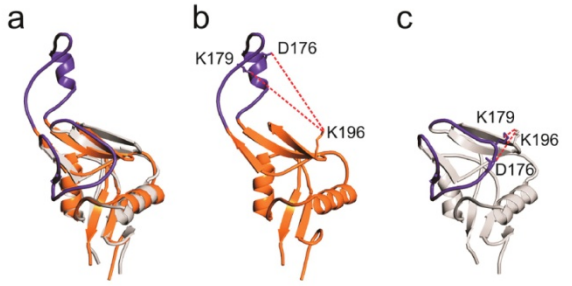


Figure 1. Cross-linking of NKR-P1A combined with mass spectrometry. (a) The different conformations of the loop region (purple) in a structural alignment of the crystal structure (orange) and the homology model (gray) of NKR-P1A. (b) The spatial distances (red dashed lines) between K179-K196 and D176-K196 in the crystal structure of NKR-P1A are inconsistent with cross-linker lengths. (c) The spatial distance constraints for K179-K196 and D176-K196 obtained in this study are consistent with the inter-residue distances in the homology model.

molecule via the domain swapping effect. Therefore, we applied mass spectrometry based techniques to investigate the dynamics and topology of this loop in NKR-P1A under near-physiological conditions.

In the first approach, NKR-P1A was subjected to chemical cross-linking.[6−8] The presence of 23 carboxyl groups and 10 primary amines in NKR-P1A provides a good basis for structural studies using this methodology. Because the high concentration of cross-linker may result in structural artifacts, it was essential to optimize molar ratio between the protein and the cross-linker. NKR-P1A was cross-linked with the zero-length cross-linker EDC (1-ethyl-3-[3-dimethylaminopropyl] carbodiimide, carboxyl-amine coupling, protein/cross-linker ratios 1:5, 1:10, 1:25, 1:100) and homobifunctional cross-linkers DSS and DSG (disuccinimidyl suberate and disuccinimidyl glutarate, amine−amine coupling, protein:cross-linker ratios 1:1, 1:3, 1:10, 1:20). Optimal results were obtained with 10× molar excess of EDC or 3× molar excess of DSS and DSG to NKR-P1A (0.2 mg/mL). Mass spectra of crude reaction mixtures showed that these conditions gave the highest relative yield of protein containing only a single intramolecular cross-link (data not shown), which minimized the likelihood of distorting the tertiary structure. The products obtained from chemical cross-linking were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and digested in gel by endoproteinases Asp-N, Glu-C, or trypsin. Proteolytic peptides were separated by a μHPLC system coupled online to an electrospray ionization Fourier transform ion cyclotron resonance (ESI FTICR) mass spectrometer. The high mass accuracy (below 2 ppm) was sufficient to provide unambiguous assignments of cross-links using MSlinks algorithm (http://ms3d.org).[6] The assignment of a cross-link was accepted only if the corresponding cross-linked peptides were generated by at least two out of three endoproteinases. Protein sequence was completely covered, and all cross-linked residues derived from the identified cross-linked peptides are listed in Table 1. For example, the signal at m/z 2506.2317 (error of 1.2 ppm) was assigned as a cross-link between lysines K166-K179 in peptides D162-S175/D176-G182 (Figure S1, Supporting Information). Out of 17 cross-linked residues identified, 6 were formed with EDC, 6 with DSG, and 5 with DSS.

Table 1. Summary of the Distance Constraints Derived from Structural Models and Comparison with the Cross-Link Distance Constraints

cross-linker	cross-linked residues	cross-link distance constraint (Å)	Cα−Cα distance (Å) crystal structure	Cα−Cα distance (Å) homology model
DSG/DSS	K166-K179	≤20 or ≤24	14.2	14.7
DSG/DSS	K179-K196	≤20 or ≤24	29.6	7.9
DSG/DSS	K148-K196	≤20 or ≤24	10.1	11.5
DSG/DSS	K146-K196	≤20 or ≤24	15.1	11.4
DSG/DSS	K125-K212	≤20 or ≤24	9.1	8.0
DSG	K146-K148	≤20	5.9	5.8
EDC	E147-K148	≤12	3.9	3.9
EDC	D121-K212	≤12	12.9	12.1
EDC	D123-K125	≤12	5.3	5.4
EDC	D143-K146	≤12	5.2	5.4
EDC	D176-K196	≤12	32.9	11.6
EDC	D176-K179	≤12	5.5	6.3

To compare the experimental inter-residue distances with the distances in the NKR-P1A crystal structure and a homology model, we used constraints based on the length of a spacer arm that is 11.4 Å for DSS, 7.7 Å for DSG, and 0 Å for EDC. Because the side chains of lysines and acidic amino acids have intrinsic flexibility, the following cutoffs are widely used: 24 Å between the α-carbons of lysines cross-linked with DSS, 20 Å with DSG, and 12 Å between the α-carbons of lysine and glutamic or aspartic acid cross-linked with EDC.[9] The Cα−Cα distances between cross-linked residues in the crystal structure are summarized and compared with Cα−Cα distances in a homology model[5] in Table 1. The inter-residue distance constraints obtained in this study are in excellent agreement with inter-residue distances in the crystal structure with the exception of K179-K196 and D176-K196 (Figure 1b). The distance between DSG cross-linked K179 from the loop region and K196 from the compact core of NKR-P1A should be ≤20 Å. Furthermore, the 32.9 Å distance between D176 from the loop region and K196 from the core in the crystal structure is inconsistent with the ability of EDC to cross-link these amino acids at a distance ≤10 Å. On the contrary, all experimental distance constraints fit the distances in the homology model of NKR-P1A (Figure 1c). However, this strategy does not reflect the steric repulsion and conformational changes in the protein backbone. In order to explore the spatial arrangement of selected residues involved in cross-linking, we performed a molecular dynamics simulation for the homology model of NKR-P1A. The minimal distance found during the 500 ns simulation between Nζ- Nζ atoms of K179-K196 is 6.1 Å, and the minimal distance between Cγ-Nζ atoms of D176-K196 is 2.6 Å. The distances predicted by molecular dynamic simulation for homology model of NKR-P1A are in agreement

with the cross-linking distance constraints and are inconsistent with the crystal structure. Taken together, the experimental data from the cross-linking study show that in solution the loop region of NKR-P1A is not extended from the compact core, as observed in the crystal structure, but is in close proximity to the core as suggested by the homology model. Further, the absence of cross-links between K179 or D176 and lysines or acidic residues, which are spread throughout the core of NKR-P1A, suggests that the backbone in the loop region is not particularly flexible.

In the second approach, H/D exchange combined with mass spectrometry was applied to study the NKR-P1A loop conformation. H/D exchange provides information on the local solvent accessibility of proteins that is valuable in studying protein dynamics, structure, and function.[10] The aim of this analysis was to compare the kinetics of H/D exchange for NKR-P1A and NKR-P1A in which the loop was removed and replaced with two alanines. This replacement was proposed by molecular dynamic simulation as the most stable mutant since the distance of the loop termini corresponds to the distance of two amino acid residues. The expression construct for nonloop NKR-P1A (NKR-P1A^{NL}) was made by crossover PCR mediated deletion−insertion mutagenesis (Figure S2, Supporting Information). NKR-P1A and NKR-P1A^{NL} were subjected to H/D exchange followed by reduction of disulfide bonds with tris(2-carboxyethyl)phosphine (TCEP). Reduced samples were loaded onto a semiautomated chromatographic system for pepsin digestion, peptide desalting (MacroTrap), and peptide separation on a reverse phased column. The extent of deuterium incorporation into each peptide was analyzed by high-resolution ESI FTICR mass spectrometry. Pepsin digestion of proteins resulted in full sequence coverage with peptic fragments approximately 5−30 residues in length (Figure S3, Supporting Information). The mass increase reflecting deuterium incorporation (Figure 2a) was followed for all identified peptides and exchange kinetics were plotted

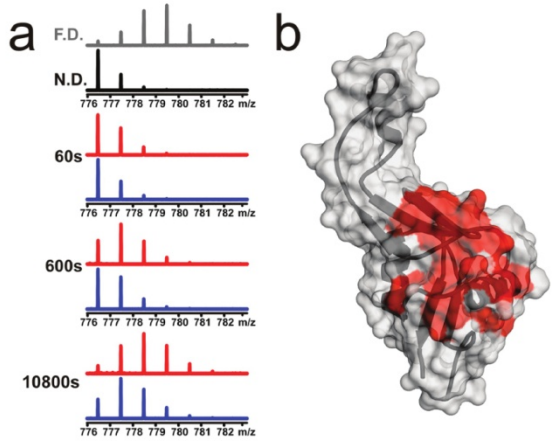


Figure 2. H/D exchange measured by mass spectrometry. (a) Changes in the isotopic pattern of doubly charged peptide 138−143 derived from NKR-P1A (blue) and NKR-P1A^{NL} (red). Three different time points are shown. N.D. represents the isotopic profile of non-deuterated peptide; F.D. represents the isotopic profile of fully deuterated peptide. (b) Differences in amide hydrogen exchange rates between NKR-P1A and NKR-P1A^{NL} mapped onto the crystal structure of NKR-P1A. The segments with a decreased exchange rate in NKR-P1A are in red.

(Figure S4, Supporting Information). Slower deuteration rates for regions 130−140, 138−143, and 191−199 were observed in NKR-P1A when compared to NKR-P1A^{NL}. This analysis suggests that these regions are protected by the loop (Figure 2b), which is consistent with the cross-linking study and the homology model but not with the crystal structure.

In order to evaluate our approach, we investigated the loop conformation of NKR-P1A by NMR spectroscopy. In the 13C-edited NOESY spectra, several cross-peaks corresponding to distances ≤5 Å between the loop region and the compact core were identified (Figure 3). Due to the unique chemical shifts of some resonances, five distances were assigned unambiguously:

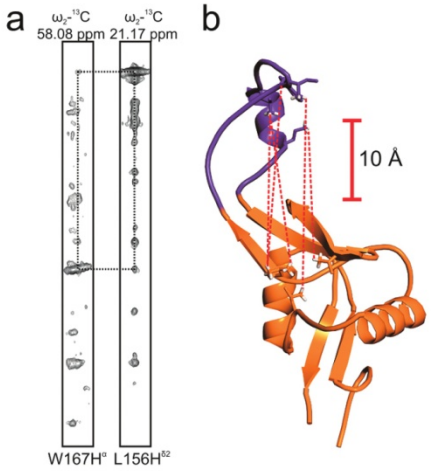


Figure 3. Conformational NMR analysis of the conserved loop of NKR-P1A. (a) Assigned cross-peaks W167Hα-L156Hδ2 and L156Hδ2- W167Hα shown in strips from a 13C-edited NOESY spectrum. (b) The distance constraints (up to 5 Å) derived from 13C-edited NOESY spectra between nearby hydrogen atoms in NKR-P1A are inconsistent with distances in the crystal structure and indicate that the loop is closely associated with compact core and association also revealed by chemical cross-linking and H/D exchange experiments.

V197Hγ2 I180Hδ1, W167Hα L156Hδ2, I168Hγ2 L129Hδ2, L173Hδ2 L156Hδ1, and L173Hδ2 L156Hδ2. These distances corroborate inter-residue distances derived from the homology model as well as data from chemical cross-linking and H/D exchange.

The approach of chemical cross-linking and H/D exchange combined with high resolution mass spectrometry described here enabled the study of structural details of the receptor NKR-P1A in solution. It revealed that the solution structure differs from the crystal structure in the conformation of the conserved loop. While the conserved loop is in close proximity to compact core in solution, it is extended from the core in the crystal structure where it interacts with the surface of a symmetry-related molecule. This suggests a testable hypothesis that the conformational change of this loop may be induced by ligand binding.

We showed that chemical cross-linking and H/D exchange coupled to high-resolution mass spectrometry provides rapid refinement of protein structure in solution. The protocol is very fast (few days), has low sample consumption (few hundreds of micrograms), and has virtually no limitations in terms of protein size.

■ ASSOCIATED CONTENT

Supporting Information

Detailed description of experimental methods for chemical cross-linking, H/D exchange, crossover PCR-mediated deletion−insertion mutagenesis, NMR analysis, and molecular dynamics calculation, integrated with figures showing representative MS analysis of cross-linked peptides, strategy of crossover mediated PCR deletion-insertion mutagenesis, sequence coverage after pepsin digestion and plots of deuterium incorporation. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: pnovak@biomed.cas.cz. Phone: +420 2 4106 2631.

■ ACKNOWLEDGMENTS

The manuscript was written through contributions of all authors. We thank Zofie Sovova and Rudiger Ettrich for helping with the design of NKR-P1A^{NL}, Frantisek Vostarek for NKR-P1A^{NL} plasmid construction, and Mark Patchett and Vladimir Havlicek for editing the manuscript. This study was supported by grants from the Grant Agency of the Czech Republic (P207/10/1040, 303/09/0477 and 305/09/H008), the Ministry of Education, Youth and Sports of the Czech Republic (LC545), the Academy of Sciences of the Czech Republic (Institutional Research Concept AV0Z50200510 and AV0Z50520701) and from the Grant Agency of Charles University (403211/2011 and Project UNCE).

■ REFERENCES

(1) Nilsson, T.; Mann, M.; Aebersold, R.; Yates, J. R.; Bairoch, A.; Bergeron, J. J. *Nat. Methods* **2010**, *7*, 681.
(2) Heck, A. J. *Nat. Methods* **2008**, *5*, 927.
(3) Vivier, E.; Tomasello, E.; Baratin, M.; Walzer, T.; Ugolini, S. *Nat. Immunol.* **2008**, *9*, 503.
(4) Kolenko, P.; Rozbesky, D.; Vanek, O.; Kopecky, V.; Hofbauerova, K.; Novak, P.; Pompach, P.; Hasek, J.; Skalova, T.; Bezouska, K.; Dohnalek, J. *J. Struct. Biol.* **2011**, *175*, 434.
(5) Sovova, Z.; Kopecky, V.; Pazderka, T.; Hofbauerova, K.; Rozbesky, D.; Vanek, O.; Bezouska, K.; Ettrich, R. *J. Mol. Model.* **2011**, *17*, 1353.
(6) Young, M. M.; Tang, N.; Hempel, J. C.; Oshiro, C. M.; Taylor, E. W.; Kuntz, I. D.; Gibson, B. W.; Dollinger, G. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5802.
(7) Rappsilber, J.; Siniossoglou, S.; Hurt, E. C.; Mann, M. *Anal. Chem.* **2000**, *72*, 267.
(8) Novak, P.; Kruppa, G. H. *Eur. J. Mass. Spectrom.* **2008**, *14*, 355.
(9) Jacobsen, R. B.; Sale, K. L.; Ayson, M. J.; Novak, P.; Hong, J.; Lane, P.; Wood, N. L.; Kruppa, G. H.; Young, M. M.; Schoeniger, J. S. *Protein Sci.* **2006**, *15*, 1303.
(10) Engen, J. R. *Anal. Chem.* **2009**, *81*, 7870.

Supporting Information

Chemical cross-linking and H/D exchange for fast refinement of protein crystal structure

Daniel Rozbesky[$,§], Petr Man[$,§], Daniel Kavan[$,§], Josef Chmelik[§], Jiri Cerny[#], Karel Bezouska[$,§] and Petr Novak[$,§,*]

[$]Department of Biochemistry, Charles University, Prague, Czech Republic

[§]Institute of Microbiology, Academy of Sciences of the Czech Republic, Prague, Czech Republic;

[#]Institute of Biotechnology, Academy of Sciences of the Czech Republic, Prague, Czech Republic.

1. Material and Methods

1.1 Protein expression and deletion of the NKR-P1A loop region

The extracellular domain of NKR-P1A was expressed in *Escherichia coli* as inclusion bodies, refolded *in vitro* by rapid dilution, and purified by ion exchange and size exclusion chromatography.[4] The non-loop NKR-P1A (NKR-P1A^{NL}) expression plasmid was constructed using crossover PCR-mediated deletion-insertion mutagenesis of the NKR-P1A expression plasmid (Figure S2). This strategy involved replacing the DNA segment encoding the loop region (T159-D187) with a 6 bp fragment encoding two alanines using the following primers: forward chimeric primer 5'-ATTGGACTAAGGTACGCAGCAAGCTGTGCTGCTA-3' and reverse chimeric primer 5'-TAGCAGCACAGCTTGCTGCGTACCTTAGTCCAAT-3'. The sequence of the inserted synthetic fragment encoding two alanines is underlined, the 5' half of the forward chimeric primer pairs with the DNA encoding I155-Y158 in the protein, and the 3' half of the forward primer encoding S188-A191. The crossover PCR deletion-insertion product was constructed in two steps. In the first step, two different PCR reactions were used to generate fragments to the left and right of the sequence targeted for deletion. The left fragment was amplified using the T7 universal and reverse chimeric primers, the right fragment was amplified using the forward chimeric and T7 terminator reverse primers. In the second step, the PCR products resulting from these two amplifications were gel-purified and the left and right fragments were annealed at their overlapping region and amplified by PCR as a single fragment using the outer T7 universal and T7 terminator reverse primers. The PCR product was ligated into the pET-

30a(+) vector and the sequence was confirmed to be correct. This plasmid was used for expression of NKR-P1A^{NL}. Protein refolding and purification was done as described for NKR-P1A.

1.2 Cross-linking of NKR-P1A

NKR-P1A protein was cross-linked at a concentration of 0.2 mg/ml with the heterobifunctional cross-linker 1-ethyl-3-[3-dimethylaminopropyl] carbodiimide hydrochloride (EDC; Pierce) and the homobifunctional cross-linkers disuccinimidyl suberate (DSS; Pierce) and disuccinimidyl glutarate (DSG; Pierce). NKR-P1A was dissolved in 50 mM pyridine-HCl buffer (pH 6.5) with 50 mM NaCl for cross-linking with EDC, and 50 mM triethylamine/bicarbonate buffer (pH 7.5) with 50 mM NaCl for cross-linking with DSS and DSG. Freshly prepared EDC solution in pyridine-HCl buffer (pH 6.5) was used at 10× molar excess to NKR-P1A, and the cross-linking reaction was allowed to proceed for 6 hours at a room temperature. Freshly prepared DSS and DSG solutions in DMSO were used at 3× molar excess to NKR-P1A and reaction mixtures were incubated for one hour at a room temperature. The duration of reactions was chosen so that all unreacted cross-linkers were hydrolyzed. The reaction in the absence of cross-linking reagent was carried out as a control. After the cross-linking reaction, the reaction mixture was subjected to SDS polyacrylamide gel electrophoresis and the band of cross-linked protein was excised. The disulfide bonds were reduced with 50 mM TCEP and free cysteines were modified with 50 mM iodoacetamide. In gel proteolysis by Asp-N, Glu-C (Roche) or trypsin (Promega) endoproteinases was carried out overnight at 37 ºC for trypsin and Asp-N and at 25 ºC for Glu-C with a enzyme/protein ratio of 1:20 (wt/wt). The peptide mixtures were desalted on a peptide MacroTrap column (Michrom Bioresources).

1.3 H/D exchange

The H/D exchange of NKR-P1A and NKR-P1A^{NL} was initiated by a 10-fold dilution in deuterated buffer (15 mM PIPES and 50 mM NaCl) at final protein concentration of 6.0 µM. The exchange was carried out at pD 6.4 and room temperature. Aliquots of 50 µl were taken after 30 s, 1 min, 3 min, 5 min, 10 min, 30 min, 1 h, 3 h and 5 h. The exchange was quenched by adding 50 µl of 0.5 M glycine-HCl buffer (pH 2.2) containing 1 M TCEP. Reduction of disulfide bonds using TCEP was allowed to proceed for 4 min at 0 ºC and samples were then frozen in liquid nitrogen. Fully deuetrated protein was prepared by incubation at 40 ºC for 36 hours.

1.4 LC-MS analysis

After chemical cross-linking, peptides were separated on a reverse phase µHPLC system (Agilent) equipped with a MAGIC C18 column (0.2 × 150 mm, Michrom Bioresources) and eluted at a flow rate of 4 µl/min under the following gradient conditions: 1-10% B in 1 min, 10-40% B in 29 min, 40-95% B in 5 min, where solvent A was 0.2% formic acid, 2.5% acetonitrile and 2.5% isopropanol in water and solvent B was 0.16% formic acid in 90% acetonitrile and 5% isopropanol.

After H/D exchange, partially and fully deuterated samples were thawed and loaded onto a semi-automated system containing injection and six port valves, a column packed with pepsin immobilized on POROS-20 AL resin (Applied

Biosystems), a peptide MacroTrap column (Michrom Bioresources) and a reverse-phase JUPITER C18 column (1.0 × 100 mm, Phenomenex). In order to minimize back-exchange, the valves, pepsin column, trap cartridge and the reverse-phase column were immersed in an ice-bath. In the first step, the protein was injected into a 100 µl loop, and then pumped through the pepsin column with solvent A containing 0.4% formic acid at a flow rate 100 µl/min. In the second step, the peptide mixture was desalted on the peptide MacroTrap column by washing with solvent A for 3.5 min and the peptides were then separated on the reverse phase column at a flow rate of 50 µl/min under the following gradient conditions: 10% C in 3.5 min, 10-45% C in 12.5 min, 100% C in 5 min, 10% C in 4 min, where solvent B consisted of 0.4% formic acid in 2% acetonitrile and solvent C consisted of 0.4% formic acid in 95% acetonitrile. The identity of generated peptides was confirmed by a separate LC-MS/MS analysis performed under the same conditions as used for LC-MS (peptide mapping). The results of peptide mapping were plotted using DrawMap script from MSTools package.[11]

1.5 Mass spectrometry and data analysis

The outlet of the reversed-phase HPLC column was connected directly to an Apex-ULTRA Qe FT-ICR mass spectrometer (Bruker Daltonics) equipped with a 9.4 T superconducting magnet via an electrospray ion source. The instrument was calibrated externally using arginine clusters resulting in mass accuracy below 2 ppm.

Data acquisition and data processing were performed using ApexControl 3.0.0 and DataAnalysis 4.0 (Bruker Daltonics), respectively. The cross-links were identified using MSlinks software (http://ms3d.org/). The MS3D algorithm was set to consider the possible single oxidation of methionine and carboxyamidomethylation of cysteines. The mass error threshold was kept below 2 ppm and all assigned fragments were verified manually.

In H/D experiments, the extracted ion chromatograms of selected peptides were plotted in DataAnalysis, signal of each peptide was averaged over the chromatographic peak and the spectrum was exported to a text file. These files were opened in mMass and the centroid values were read out. [12] Corrected percentages of deuteration were calculated according to this equation: $corr\ \%D = 100*[(M_D-M_N)/(M_F-M_N)]$ where M_D is average mass (centroid value) of partially deuterated peptide, M_F average mass of fully deuterated peptide, and M_N average mass of non-deuterated peptide. Plots of deuteration percentage versus time were drawn using MS Tools.[11]

1.6 NMR spectroscopy

For NMR experiments, ^{15}N-^{13}C-NKR-P1A was expressed in *E. coli* grown in M9 minimal medium with $^{15}NH_4Cl$ as a source of nitrogen and ^{13}C-glucose as a source of carbon. The NMR samples consisted of 0.6 mM ^{15}N-^{13}C-NKR-P1A in 15 mM PIPES buffer (pH 6.8), 50 mM sodium chloride, 1 mM sodium azide and 10% deuterium oxide. All NMR spectra were recorded at 20 °C on a Bruker Avance III 600 MHz spectrometer equipped with a cryogenic 1H/13C/15N TCI probe head. Backbone assignments were obtained from standard triple-resonance experiments, including HNCACB, CBCA(CO)NH, HN(CA)CO, and HNCO. Side-chain assignments were based on the HCCH-

TOCSY spectrum. The ^{13}C and ^{15}N-edited NOESY spectra were recorded using a mixing time of 120 ms. All spectra were processed using NMRPipe[13] and analyzed with Sparky software.[14]

1.7 Molecular dynamics simulation

The molecular dynamics simulation was performed using OpenMM[15] and Zephyr[16] implementing the GPU accelerated version of the GROMACS[17] suite. Implicit solvation (GBSA, ε=78.3, "accurate water" with collision interval of 10.99 fs) in combination with the parm96 force-field[18] was used. The initial structure was optimized and followed by simulation at 300 K with a time step of 2 fs. Snapshots of the geometry were saved every 10 ps throughout the 500 ns simulation. The resulting trajectory was analyzed using the g_dist program of GROMACS, measuring the distances between C atoms of selected residues. The whole trajectory was further analyzed by the g_rmsf program of GROMACS to obtain the atom fluctuations within the protein.

2. Supporting references

(11) Kavan, D., Man, P. *Int. J. Mass. Spectrom.* **2011**, *302, 53.*

(12) Strohalm, M.; Kavan, D.; Novak, P.; Volny, M.; Havlicek, V. *Anal. Chem*. **2010,** *82, 4648*.

(13) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR.* **1995**, *6, 277.*

(14) Goddard, T. D.; Kneller, D. G. *SPARKY 3*, University of California, San Francisco.

(15) Eastman, P.; Pande V. S. *Comput. Sci. Eng.* **2010,** *12, 34*.

(16) Friedrichs, M. S.; Eastman, P.; Vaidyanathan, V.; Houston, M.; Legrand, S.; Beberg, A. L.; Ensign, D. L.; Bruns, C. M.; Pande V. S. *J. Comp. Chem.* **2009**, *30, 864*.

(17) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J. Comput. Chem.* **2005,** *26, 1701.*

(18) Kollman, P.A. *Acc. Chem. Res.* **1996,** *29, 461.*

3. Supporting figures

Figure S1. Identification of cross-links. **(a)** Representative base-peak ion chromatogram from the LC-MS analysis of Asp-N digested DSG-cross-linked NKR-P1A. **(b)** MS spectrum from the LC-MS analysis in (a). The cross-link was identified using MSlinks algorithm (http://ms3d.org).[6] The triply-charged molecule ($[M+H]^+$= m/z 2506.2317) corresponds to cross-linked peptides between lysines K166-K179.

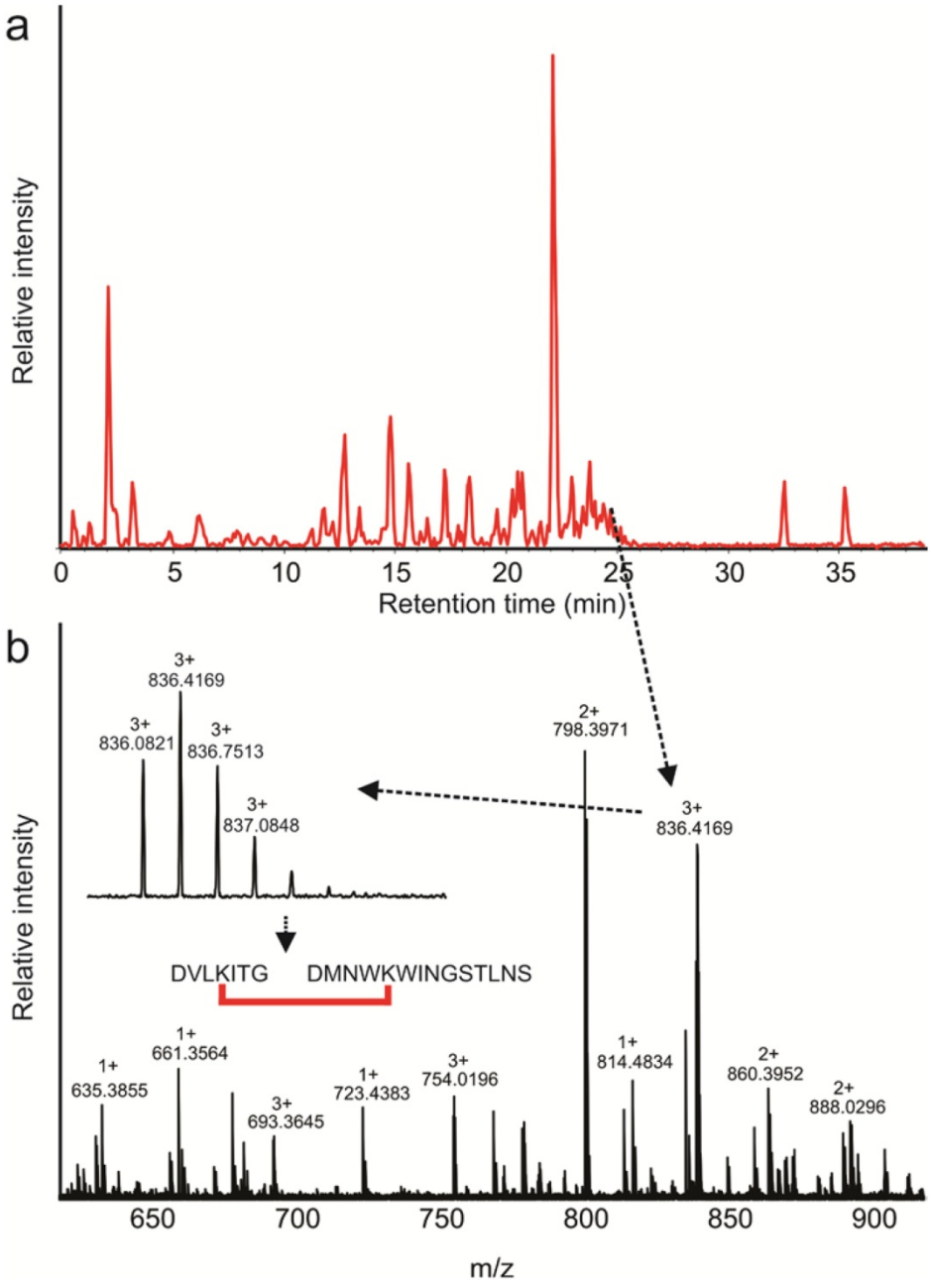


Figure S2. Strategy for constructing the NKR-P1A^{NL} expression plasmid using crossover PCR mediated deletion-insertion mutagenesis. The top circle represents an expression plasmid containing a gene of NKR-P1A (orange) with the sequence encoding the conserved loop (red) subcloned into the pET-30a(+) expression vector (gray). Two PCRs were used to generate two fragments which formed an in-frame deletion of the segment encoding the loop region when fused. The segment encoding the loop region was replaced by a fragment encoding two alanines (blue).

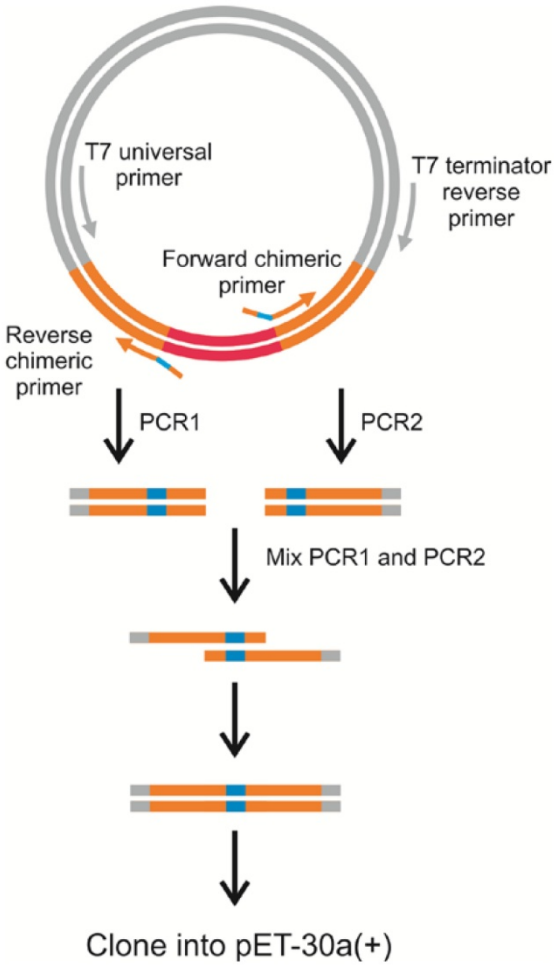


Figure S3. Peptide mapping of NKR-P1A and NKR-P1A^{NL} after digestion with immobilized pepsin. All peptides identified by ESI-FT ICR-MS/MS analysis are shown as blue (NKR-P1A) and red (NKR-P1A^{NL}) bars. Residue numbering was identical for NKR-P1A and NKR-P1A^{NL} numbering was kept after loop deletion unchanged with the gap 158-188.

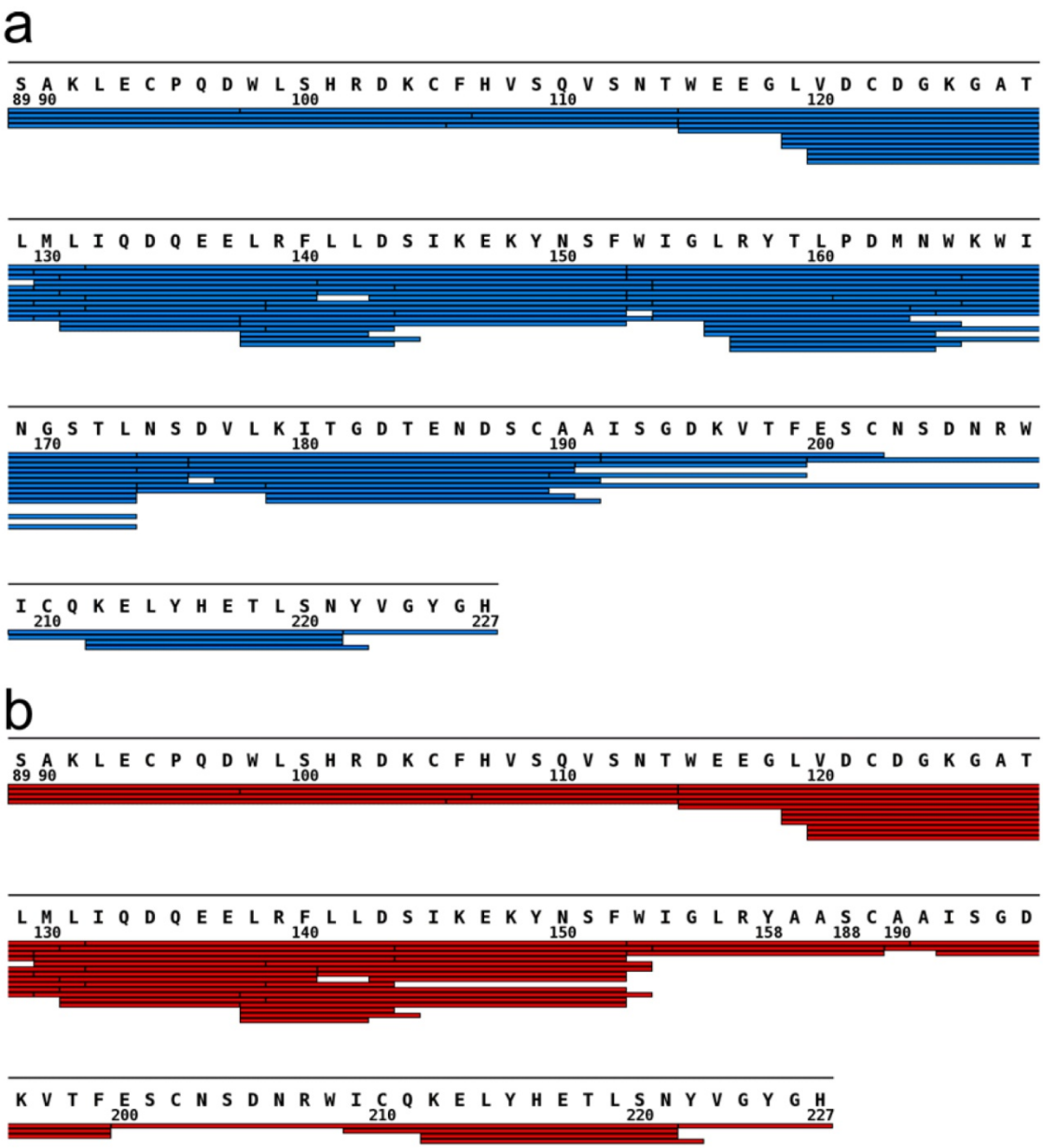
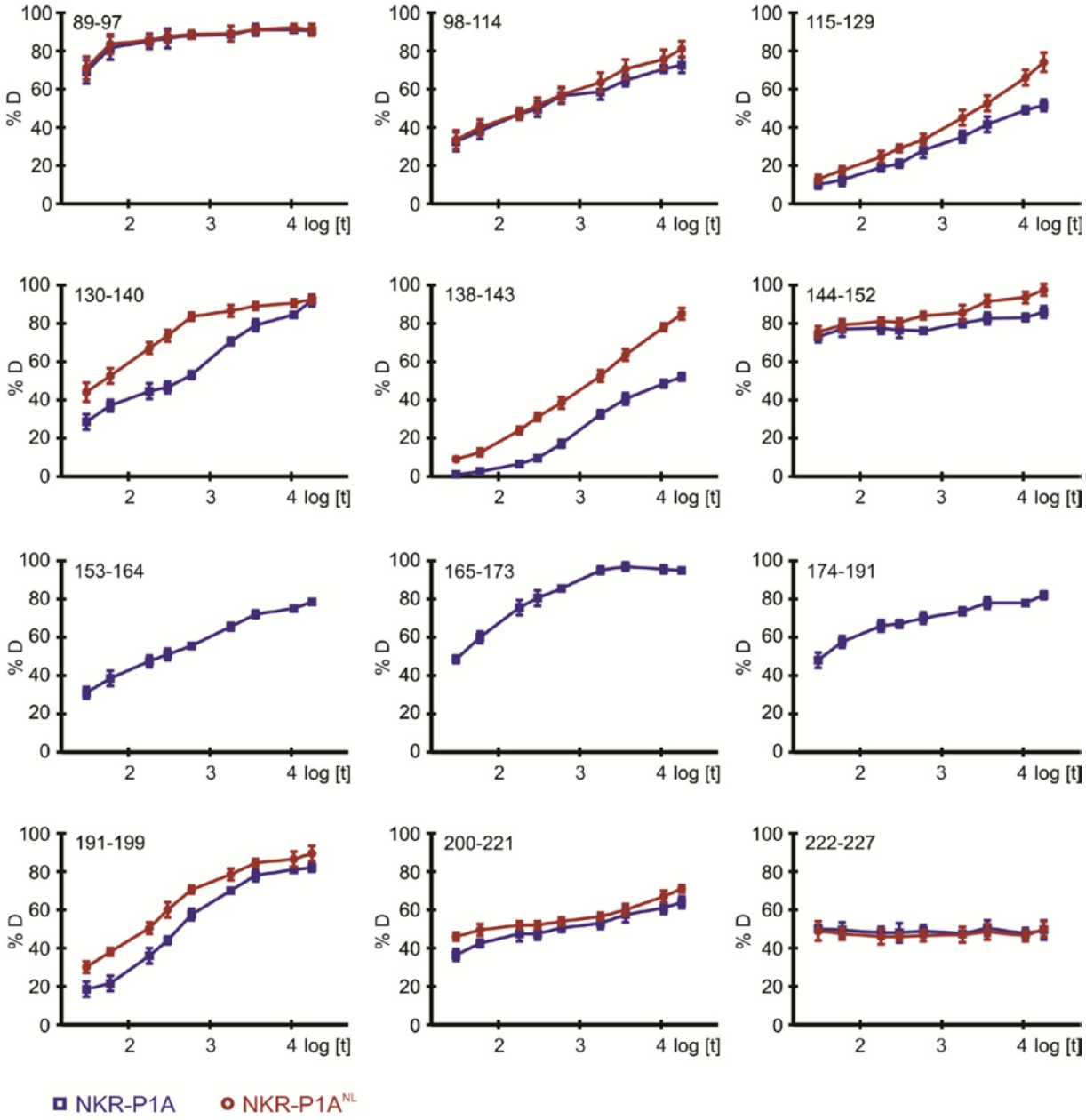


Figure S4. Time profiles of deuterium incorporation for representative peptides of NKR-P1A (blue) and NKR-P1A^{NL} (red). The plots show percentage deuteration versus time. The error bars represent the standard deviations for the two independent determinations of the amide hydrogen exchange.



PAPER VI

Rozbesky, D., Sovova, Z., Marcoux, J., Man, P., Ettrich, R., Robinson, C.V., Novak, P.

Structural model of lymphocyte receptor NKR-P1C revealed by mass spectrometry and molecular modeling.

Anal. Chem. **85**, 1597-604 (2013).

My contribution to the publication: *research design, research performing (protein expression, refolding in vitro, purification, chemical cross-linking, disulfide mapping), data collection, data analysis and interpretation, manuscript writing*

analytical chemistry

Article

pubs.acs.org/ac

Structural Model of Lymphocyte Receptor NKR-P1C Revealed by Mass Spectrometry and Molecular Modeling

Daniel Rozbesky,[†,‡] Zofie Sovova,[§,∥] Julien Marcoux,[⊥] Petr Man,[†,‡] Rudiger Ettrich,[§,∥] Carol V. Robinson,[⊥] and Petr Novak*[,†,‡]

[†]Department of Biochemistry, Charles University, Prague, Czech Republic

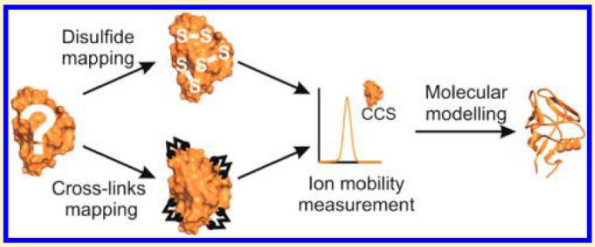
[‡]Institute of Microbiology, Academy of Sciences of the Czech Republic, Prague, Czech Republic

[§]Institute of Nanobiology and Structural Biology, Global Change Research Center, Academy of Sciences of the Czech Republic, Nove Hrady, Czech Republic

[∥]Faculty of Sciences, University of South Bohemia, Nove Hrady, Czech Republic

[⊥]Department of Chemistry, University of Oxford, Oxford, U.K.

ABSTRACT: NKR-P1C is an activating immune receptor expressed on the surface of mouse natural killer cells. It has been widely used as a marker for NK cell identification in different mice strains. Recently we solved a crystal structure of the C-type lectin-like domain of a homologous protein, NKR-P1A, using X-ray crystallography and also described the strategy for rapid characterization of the protein conformation in solution. This procedure utilized chemical cross-linking, hydrogen/deuterium exchange, and molecular modeling. It was found that the solution structure differs from the crystal structure in the conformation of the loop region. The loop, detached from the protein compact core in the crystal structure, is closely attached to the core of the protein in solution. Here we present and interpret the solution structure of the C-type lectin-like domain of NKR-P1C using chemical cross-linking and molecular modeling. The validation of the model and conformation of the loop region in NKR-P1C were addressed using ion-mobility mass spectrometry.



The C-type lectin-like receptor NKR-P1C is an activating immune receptor expressed on the surface of natural killer (NK) cells, key players in antiviral and antitumor immunity.[1] NKR-P1C, originally known as NK1.1 antigen, is a common marker for the identification of NK cells in certain mice strains.[2,3] It also plays a crucial role in the stimulation of NK cell cytotoxicity of target cells and appears to be involved in natural killing of some tumor targets.[4,5] Antibody cross-linking of NKR-P1C receptor leads to Ca^{2+} flux, phosphatidylinositol turnover, kinase activity, and cytokine production. It was also reported that its intracellular tail associates with the tyrosine kinase, p56 (lck).[6] From a structural point of view, the NKR-P1C is a type II transmembrane protein, and in the literature it is described as a disulfide-linked homodimer with a fold typical for the C-type lectin-like family.[7,8] Although the NKR-P1C was discovered more than 20 years ago, its structure remains unknown. Interestingly, NKR-P1 family includes both activating and inhibitory receptors and the C-type lectin-like domain of NKR-P1 family members share considerable homology; the structural origins of ligand-binding specificity are thus of high interest.[9]

Structure of the extracellular part, namely, the C-type lectin-like domain, of a homologous protein, NKR-P1A, was solved recently using X-ray crystallography. However, the fold of the protein found in the crystal differed from the fold of known homologous proteins (e.g., CD69, NKG2D) in the loop region which was detached from the protein compact core and provided significant interaction surface with the other molecules of NKR-P1A in the crystal.[10] In order to validate the structure in solution and at near physiological conditions, we developed a protocol combining chemical cross-linking, deletion of the loop region, hydrogen/deuterium exchange coupled to mass spectrometry, and homology modeling. Surprisingly we found that the loop region was collapsed and closely interacting with the rest of the protein.[11] Sequence and phylogenetic analysis suggests that an extended loop region evolved in concert with target specificity, adapting the chemical properties of the residues to specific target ligands while embedding them in a structurally stable framework, suggesting that the long loop plays a key role in ligand specificity, and that a conformational equilibrium might be involved in ligand binding.[9]

In this study we use a set of methods, ranging from chemical cross-linking with sequence threading using mass spectrometry, restraint-based molecular modeling and steered molecular dynamics, to ion-mobility mass spectrometry, not only to

Received: October 2, 2012
Accepted: December 18, 2012
Published: December 18, 2012

describe the overall fold of NKR-P1C but also to elucidate the solution structure of the large extended loop described above.

Traditional methods of protein structure determination rely on high-resolution techniques: X-ray crystallography and NMR spectroscopy. However, with the technological advances and increasing demands to study protein structure of dynamic complexes under physiological conditions,[12] complementary low-resolution methods are becoming more widely used. The main advantages of low resolution techniques are the low sample requirement, which applies to proteins difficult to express/isolate, and the ability to work at near-physiological conditions with dynamic, heterogeneous proteins at low concentrations. Moreover, there is virtually no limitation in terms of protein size, flexibility, and time requirements for both data acquisition and data processing are relatively low. Low-resolution approaches usually apply several methodological concepts in concert, and their main role is in characterization of proteins or protein systems not amenable by NMR or X-ray.[13] An excellent example of this approach is the description of the interaction between the 14-3-3 protein and the regulator of G-protein signaling (with an unstructured N-terminal half) by combination of small-angle X-ray scattering, spectroscopic techniques, hydrogen/deuterium exchange coupled to mass spectrometry and molecular modeling.[14] Other examples include description of quaternary structure of eukaryotic RNAPII complex or assembly of proteasome.[15,16]

The concept of chemical cross-linking in combination with sequence threading was introduced more than a decade ago to determine the fold of bovine basic fibroblast growth factor 2 (FGF-2).[17] Since then, the implementation of protein cross-linking has expanded and various problems were addressed using this technique, protein conformational changes,[18] details of protein structure,[11] a mechanism of small proteins,[19] a topology of protein−peptide complexes,[20] or an architecture of multiprotein complexes.[21,22] In chemical cross-linking, side chains of amino acid residues are targeted by bifunctional reagents in which the two reactive groups are interconnected by a spacer arm of variable length. If two residues are within a certain distance from each other, they may be covalently connected by a cross-linking reagent. In this approach, mass spectrometry is used for identification of cross-linked peptides and also for exact localization of the cross-link.[23,24] Identification of cross-linked peptides may provide two types of information. One is simple assignment of protein regions/residues involved in the interaction. This is the main aim of large scale studies mapping the entire cellular interactome.[25] A second way to use information from chemical cross-linking is to derive distance constraints in order to generate low resolution structure. Here, the number of cross-links (distances) dictates the resolution of the final structure. High-resolution structure requires $3N$ distances where N is the number of amino acid residues within a protein. However, low-resolution structure or fold identification can be done with less information, and $N/10$ distances were described as a minimal data set.[17] Therefore, a variety of cross-linkers must be employed to overcome possible problems such as low reactivity[26] and to generate enough constraints.[18,27] Taken from this point of view, disulfide bridges stabilizing the protein structure[28,29] represent a native cross-link and their identification may facilitate protein homology modeling and quality control in in vitro recombinant protein production.[30,31] Tools for disulfide pattern analysis have advanced significantly in the past 2 decades. Mass spectrometric-based strategies of disulfide arrangement are based on either partial reduction of a protein in conjunction with cysteine alkylation[32] or hydrolysis of intact proteins and subsequent analysis of disulfide-linked peptides[33,34] or top down strategies.[35]

In cases where at least one member of a "fold-family" is known, homology modeling can be used to analyze the three-dimensional structure of proteins of unknown structure. The approach chosen in this study is a back and forth between experiment and theory and represents a cycle of structure generation and verification, in which distance constraints derived from disulfides and chemical cross-linking of NKR-P1C are not only directly implemented in the structure modeling but also used in the next cycle for the experimental validation of the proposed structure and the improvement of the modeling accuracy. However, because of differences between the crystal structure and in-solution structure of NKR-P1A mentioned above, we put the strong focus on the loop position also in NKR-P1C. H/D exchange and deletion of the loop was used previously to assess the loop conformation in NKR-P1A. The results were consistent with the output from the chemical cross-linking but instead of deleting a large portion of the protein we used here native electrospray coupled to ion mobility mass spectrometry (IM-MS) to differentiate between the two possible loop conformations.

This technique can be seen as gas-phase chromatography, in which the ionized molecules are transmitted in a drift cell where they are submitted to a weak electric field and a relatively high pressure of neutral gas (typically helium or nitrogen) to separate them on the basis of their collisional cross section (CCS). Therefore, even the species with the same molecular weight but different conformation can be effectively separated. Since its first applications to small proteins in the early 90s,[36] IM-MS has progressed significantly, enabling nowadays the study of systems of increasing complexity. Since the work of Ruotolo et al.[37] confirming that quaternary structures could be maintained in the gas phase, IM-MS has provided additional information about large protein complexes, especially when combined with other structural techniques such as EM,[38] NMR,[39] or cross-linking.[16] This hybrid approach was shown to be invaluable for model filtering[40] or for the analysis of conformational isoforms.[41,42]

■ EXPERIMENTAL SECTION

Protein Expression and Purification. The extracellular portion of the C57BL/6 mouse NKR-P1C (residues 89−223) was produced in *E. coli* into inclusion bodies. Protein was refolded in vitro and purified using ion-exchange chromatography as described previously.[43] The purified protein was dialyzed against buffer containing 10 mM HEPES (pH 7.0), 150 mM NaCl, and 1 mM NaN_3, and protein was concentrated to 10 mg/mL using Amicon Ultra centrifugal filter unit (Millipore). The monoisotopic mass of intact protein was measured on a 9.4 T Apex Ultra Qe Fourier transform ion cyclotron resonance (FTICR) mass spectrometer (Bruker Daltonics), confirming the identity of the protein.

Disulfide Bonds Determination. The disulfide bonds in NKR-P1C were determined according to the previously published protocol.[34] Briefly, the protein was separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and digested by trypsin or Asp-N protease under nonreducing conditions in the presence of 200 μM cystamine. The peptide mixtures were desalted on a peptide MacroTrap (Michrom Bioresources).

Chemical Cross-Linking. The NKR-P1C protein was cross-linked at a concentration of 0.2 mg/mL with the homobifunctional cross-linkers disuccinimydyl suberate (DSS) and disuccinimydyl glutarate (DSG). For the cross-linking reaction, NKR-P1C was dialyzed into 50 mM triethylamine/bicarbonate buffer (pH 7.5) with 50 mM NaCl, and cross-linkers DSS and DSG were used as 1:1 (mol/mol) mixtures of nondeuterated and four-times deuterated compounds (d_0/d_4). Freshly prepared stock solutions of cross-linkers (10 mg/mL in DMSO) were added in 10× molar excess to NKR-P1C, and reaction mixtures were incubated for 1 h at room temperature.

Electrophoretic Separation and In Gel Digestion. Following cross-linking, NKR-P1C was separated on a Nu-PAGE 4−12% Bis-Tris gel using MES running buffer and the bands corresponding to monomeric cross-linked protein were excised. The disulfide bonds were reduced with 50 mM TCEP for 10 min at 90 °C, and free cysteines were alkylated with 50 mM iodoacetamide for 90 min at room temperature in the dark. Trypsin and Asp-N digestion proceeded overnight at 37 °C with an enzyme/protein ratio of 1:20 (w/w). The resulting peptide mixtures were desalted on a peptide MacroTrap column (Michrom Bioresources).

LC−MS Analysis. After desalting, the samples (disulfide bond assignment or chemical cross-linking) were loaded onto a reverse phased column MAGIC C18 column (0.2 mm × 150 mm, Michrom Bioresources) and separated on a capillary HPLC system (Agilent Technologies) at a flow rate of 4 μL/min under the following gradient conditions: 1−10% B in 1 min, 10−40% B in 29 min, 40−95% B in 5 min, where solvent A was 0.2% formic acid, 2.5% acetonitrile, and 2.5% isopropanol in water and solvent B was 0.16% formic acid in 90% acetonitrile and 5% isopropanol. The column was connected directly to an Apex-ULTRA Qe FTICR mass spectrometer (Bruker Daltonics) equipped with a 9.4 T superconducting magnet using an electrospray ion source. The instrument was calibrated externally using arginine clusters resulting in mass accuracy below 2 ppm.

Data Analysis. Data acquisition and data processing were performed using ApexControl 3.0.0 and DataAnalysis 4.0 (Bruker Daltonics), respectively. The disulfide bonds and cross-links were identified using Links software.[17] The Links algorithm was set to consider the possible single oxidation of methionine, and different cysteine modifications were set for disulfide mapping (disulfide bond formation) and for cross-linking (carbamidomethylation). The enzyme specificities were set to K-/R- for trypsin and -D/-E for Asp-N and the missed cleavages were set to 3. Modification or cross-link of amino groups (protein N-terminus and lysine side chains) by DSS or DSG was considered for cross-linking experiments. The mass error threshold was kept below 3 ppm, and all assigned fragments were verified manually.

Homology Modeling and Molecular Dynamics. Homology modeling followed by a short steepest descent minimization was performed using the MODELER 9v7 package.[44] The recently published crystal structure of mouse NKR-P1A[10] (pdb-code 3M9Z) was used as a template for the core of the protein, and the published structure of human CLEC5A[45] (pdb-code 2YHF) was used for the extended loop region. This structure was chosen as it has the highest homology with our sequence (35%) in the extended loop region. In total, 10 models were generated and the best one was chosen based on the distribution of the torsional angles of amino acids in favored regions in the Ramachandran diagram and on visual inspection of the model (both in Swiss PDBViewer[46] and the MODELER objective function). The three-dimensional model is based on all non-hydrogen atoms. All three disulfide bridges were already created during the restraint-based modeling procedure and refinement was achieved through algorithmic analysis and minimization with the Amber03 force field56 in YASARA[47] after hydrogen atoms were added. Confrontation of the best "blind" model with cross-linking data showed that all distances determined experimentally are in agreement with the model. Constraints for the loop region coming from the experimentally determined distances were solved by nonequilibrium molecular dynamics simulation in vacuo using YASARA with the Amber03 force field applying an external force on the loop atoms in the direction of coordinates within the experimentally determined range. A nonbonded interaction cutoff of 7.8 Å was applied using a multiple time step of 1 fs for intramolecular and 2 fs for intermolecular forces and the simulation was then run at 298 K and constant pressure. The external forces were interactively applied to the loop unless the cross-linking distances were fulfilled. During the nonequilibrium molecular dynamics, the rest of the system was fixed to prevent distortion of the protein core. Additionally, torsion angles in K212 were manually adjusted to reorient the side chain in the opposite direction. Once all experimental constraints were fulfilled, the whole structure was minimized by 2000 steps of simulated annealing minimization.

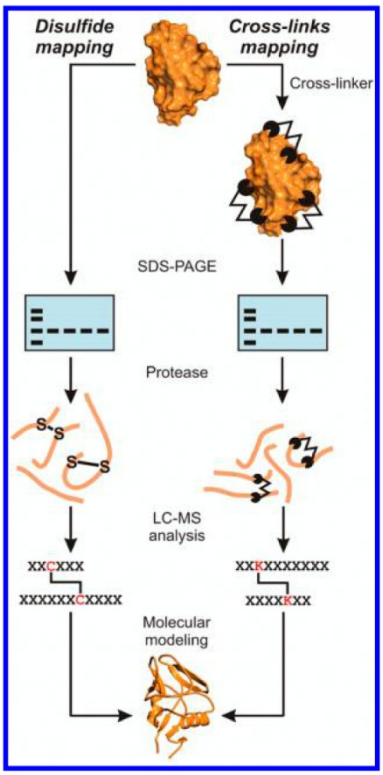
Native Mass Spectrometry and Ion Mobility Measurement. The protein was buffer exchanged in 200 mM ammonium acetate pH 7.4 using a micro BioSpin-6 device (Biorad) and diluted in the same buffer to reach 13 μM concentration. The sample was then loaded into gold-coated nanoflow capillaries prepared in house.[48] IM-MS data was acquired on a quadrupole ion-mobility time-of-flight mass spectrometer (Synapt HDMS, Waters) modified as described previously such as the collisional cross section (CCS) can be determined directly by plotting the drift time versus the reciprocal drift voltage[40,49] as long as the temperature and pressure in the IM cell are determined accurately. The capillary, cone, and trap collision energy were set at 1.6 kV, 20 V, and 10 V, respectively, in order to minimize unfolding in the gas phase. The backing pressure was set at 6−8 mbar, and the pressure in the drift cell was set at 2.17 Torr using a flow of helium of 50 mL/min. The temperature outside the drift cell was monitored with a thermocouple (Omega Engineering). The drift voltage was increased from 50 to 200 V in 5−30 V steps. The drift time for each charge state and each drift voltage were obtained using MassLynx and DriftScope (Waters). Calibration was performed using a 10 mg/mL solution of cesium iodide. Theoretical CCS for the NKR-P1C and its homologue NKR-P1A (pdb code 3M9Z) were calculated using the projection approximation (PA) method implemented in MOBCAL and scaled, as described previously.[50]

■ RESULTS AND DISCUSSION

Disulfide Mapping of Mouse NKR-P1C. The C-type lectin-like domain of recombinant mouse NKR-P1C (NKR-P1C) contains 135 amino acid residues including six cysteine residues. In order to determine the cysteine status, we measured the mass of the intact protein by electrospray (ESI)-FTICR MS. Three disulfide bonds seemed to be present in the NKR-P1C as the monoisotopic mass of the intact protein $[M + H]^+$ of 15 394.4421 Da was in agreement with the predicted monoisotopic mass calculated with the assumption of

all paired cysteine residues. The experimental setup for the disulfide mapping is depicted in Scheme 1. Digestion of the protein by Asp-N or trypsin proteases generated peptides containing disulfide-linkages that were then subjected to reverse phase capillary HPLC followed by ESI-FTICR MS. After Asp-N digestion, an unambiguous disulfide pattern of NKR-P1C was determined. The first disulfide bond between Cys94 and Cys105 was assigned based on the dipeptide with the monoisotopic mass $[M + H]^+$ of 2951.2893 (Figure 1a). The other cystine dipeptide with the monoisotopic mass $[M + H]^+$ of 3468.7075 (Figure 1b) indicated the presence of a bond between Cys122 and Cys 210. Finally, the monoisotopic mass $[M + H]^+$ of 2910.4065 (Figure 1c) matches the dipeptide with a link between Cys189 and Cys202. Trypsin digestion provided verification of the assigned pattern by identification of a cystine dipeptide with the monoisotopic mass $[M + H]^+$ of 2840.3059 (disulfide bond between Cys189 and Cys202) and tripeptide containing the four remaining cysteines, all involved in disulfide bonds (monoisotopic mass $[M + H]^+$ 4859.1958). All experimentally determined disulfide linkages in NKR-P1C are summarized in Table 1 and in Figure 1d.

Scheme 1. Schematic Workflow of Mass Spectrometric Disulfide and Cross-Link Mapping[a]



[a]In the analysis of disulfide pattern, the protein is separated using SDS-PAGE containing cystamine to prevent disulfide bonds from scrambling. Protein is subsequently in-gel digested, and peptides are analyzed by capillary liquid chromatography coupled to high-resolution mass spectrometry. Accurate masses of intact peptides and disulfide linked peptides are subjected to database searching to identify disulfide-linked peptides. In chemical cross-linking, the protein is cross-linked in solution with a mixture of light and heavy cross-linkers and cross-linked protein is separated using SDS-PAGE. Cross-links are identified in the same fashion as disulfide-linked peptides. Finally, disulfide pattern and distance constraints derived from cross-links are used in homology modeling.

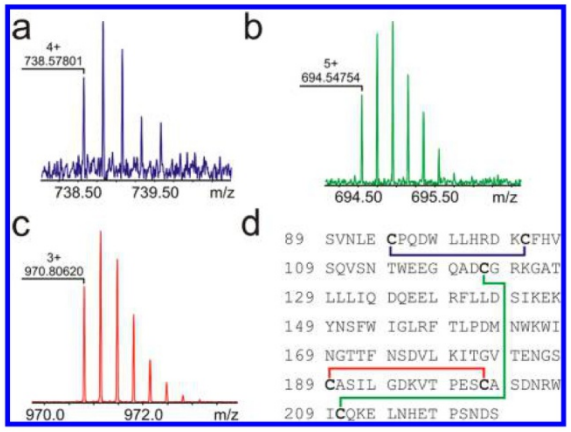


Figure 1. Disulfide mapping of mouse NKR-P1C. (a) Representative mass spectra of cystine peptide Cys94/Cys105, (b) cystine peptide Cys122/Cys210, and (c) cystine peptide Cys189/Cys202 from Asp-N digest. (d) Protein sequence of the C-type lectin-like domain of NKR-P1C with schematic drawing of disulfide arrangement.

Table 1. Summary of Disulfide Bond Analysis by LC–MS[a]

Cys-Cys cross-link	disulfide-linked peptides(s)	theoretical $[M + H]^+$	experimental $[M + H]^+$	error (ppm)
Cys94/Cys105[a]	89-96/103-120	2951.2934	2951.2893	1.5
Cys122/Cys210[a]	121-133/205-221	3468.7111	3468.7075	1.2
Cys189/Cys202[a]	176-194/195-204	2910.4071	2910.4065	0.3
Cys94/Cys105[t]	89-124/208-212	4859.1973	4859.1958	0.3
Cys122/Cys210[t]	89-124/208-212	4859.1973	4859.1958	0.3
Cys189/Cys202[t]	180-196/197-207	2840.3037	2840.3059	0.6

[a]Identified cross-linked peptides and corresponding disulfide bonds are shown in the first two columns. Upper indices [a] and [t] in the first column indicate the protease (a, AspN; t, trypsin) by which the protein was digested. Theoretical and experimental masses and mass errors are shown in the third and fourth columns, respectively.

Chemical Cross-Linking and Mass Spectrometry. The presence of nine lysine residues in NKR-P1C provides a good basis for structural studies using chemical cross-linking of primary amines. The schematic workflow of the cross-linking and cross-links identification is depicted in Scheme 1. The NKR-P1C was incubated with 10× molar excess of DSS or DSG. Both cross-linkers were applied as a 1:1 (mol/mol) mixture of the nondeuterated and deuterated (d_0/d_4) derivate to then facilitate the identification of cross-linked peptides. The concentration of cross-linking reagents was optimized to give the highest relative yield of protein containing a single intramolecular cross-link, which minimized the likelihood of distorting the tertiary structure. The reactions were not quenched but the cross-linkers were allowed to hydrolyze in order to reduce the complexity of the reaction mixture. Under these conditions, no aggregates were observed on SDS-PAGE and position of the cross-linked NKR-P1C corresponded to the

control represented by the unmodified protein (Figure 2a). Following the in-gel reduction and alkylation of cysteines, the

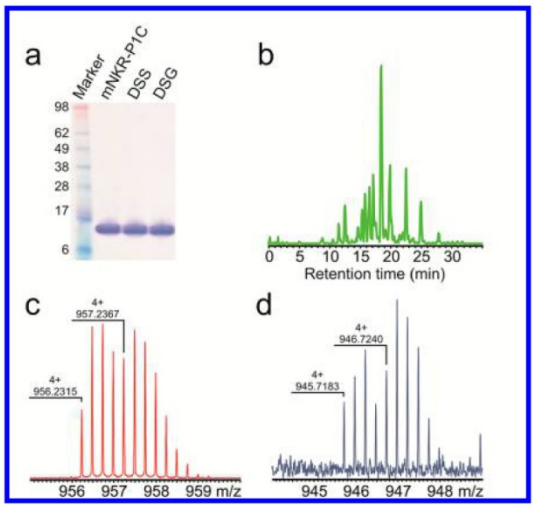


Figure 2. Chemical cross-linking and mass spectrometric cross-links identification. (a) Coomassie stained SDS PAGE gel of unmodified NKR-P1C which was used as a sample control, and the NKR-P1C cross-linked with DSS-d_0/d_4 and DSG-d_0/d_4. (b) Representative total ion chromatogram of the reverse-phase capillary HPLC and FTICR MS of trypsin-digested NKR-P1C cross-linked with DSS. (c, d) Representative mass spectra of peptides 125-139/208-223 cross-linked with DSS-d_0/d_4 or DSG-d_0/d_4. Because the protein was cross-linked using a 1:1 mixture of light and heavy cross-linker, the peptides containing the cross-linker are recognized by their distinct doublet isotope pattern with a mass difference of 4.025 Da in the deconvoluted mass spectra.

NKR-P1C was digested by trypsin or Asp-N. Generated peptides were separated by capillary HPLC connected to ESI-FTICR MS (Figure 2b). Although the peptides with deuterated cross-linkers showed small retention time shift compared to their nondeuterated counterparts, the difference did not complicate the data analysis. The cross-links were identified using Links software,[17] which matches a peak list obtained from the LC−MS analysis with theoretical masses of all possible products. The use of heavy and light cross-linkers resulted in characteristic isotopic pattern of peptides bearing the cross-linker (mass difference of 4.025 Da in the deconvoluted mass spectra, Figure 2c,d). This specific signature, together with high mass accuracy (below 2 ppm) provided unambiguous assignment of the cross-links. All unique cross-linked residues derived from the identified cross-links are listed in Table 2. Out of 10 cross-linked residues identified, 7 were formed with DSS and 3 with DSG. Similar cross-linking pattern was observed using DSS and DSG, with the exception of Lys146-Lys148, Lys148-Lys196, and Lys104-Lys125, which were not detected with the shorter cross-linker (DSG). All experiments including disulfide bond assignment as well as chemical cross-linking were done in triplicate and all presented cross-links were identified in all three repetitions. There were no cross-links found in just one or two experiments confirming the overall structural homogeneity of the protein.

Molecular Modeling. The three-dimensional model of NKR-P1C shown in Figure 3a has the typical C-type lectin-like fold and three disulfide bridges. Within the whole cross-linking network (Figure 3b) only Lys166 and Lys179 are part of the

Table 2. Table of Identified Cross-Links[a]

cross-linking reagent	cross-linked peptide(s)	Lys-Lys cross-link	theoretical d_0-$[M + H]^+$	experimental d_0-$[M + H]^+$	error (ppm)
DSS	103-120/121-133[a]	K104-K125	3703.7847	3703.7915	1.8
DSS	125-139/208-223[t]	K125-K212	3821.9015	3821.9043	0.6
DSS	140-157[t]	K146-K148	2367.2747	2367.2663	2.7
DSS	147-157/180-207[t]	K148-K196	4488.1446	4488.1382	1.5
DSS	162-175/195-204[a]	K166-K196	2944.3142	2944.3142	0.3
DSS	176-194/195-204[a]	K179-K196	3164.5338	3164.5356	0.5
DSS	186-200/201-213[a]	K196-K212	3338.5059	3338.5059	0.9
DSG	125-139/208-223[t]	K125-K212	3779.8545	3779.8511	1.0
DSG	162-175/195-199[a]	K166-K196	2368.0914	2368.0859	2.5
DSG	176-194/195-204[a]	K179-K196	3122.4868	3122.4951	2.5

[a]First column indicates the cross-linking reagent used, second column shows the peptide/dipeptide based on which the cross-link (third column) was identified. Type of protease used for digestion of the cross-linked protein is shown as the upper indices in the second column: a, AspN; t, trypsin. Theoretical and experimental masses of nondeuterated cross-link and mass errors are shown in the fourth, fifth, and sixth columns, respectively.

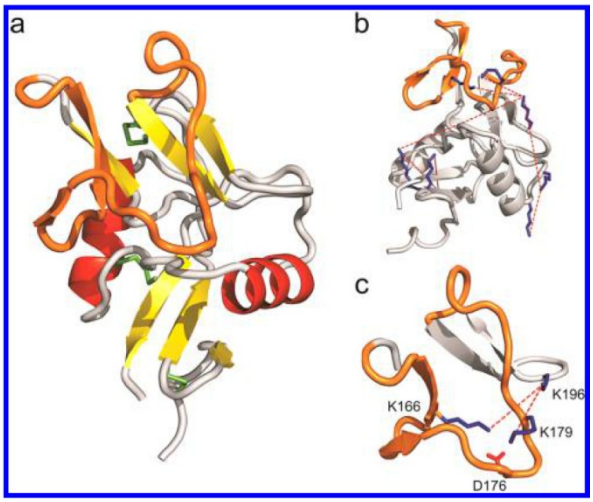


Figure 3. (a) Structural model of the mouse NKR-P1C with the loop region colored in orange and disulfide bonds colored in green. (b) Model of NKR-P1C with lysine residues involved in chemical cross-linking. Cross-linked residues are highlighted using dashed line. (c) Detail view on the loop region with K166 and K179 interacting with the negatively charged aspartic acid D176. These two residues were shown to be cross-linked to K196.

loop region, whereas the other lysine residues are located in the globular domain. The experimental constraints matched conformations in which the loop is packed onto the protein surface (cross-links between Lys196 from the core region and Lys166 or Lys179 from the loop region) in a similar way as in the dengue virus receptor CLEC5A.[45] Furthermore, these distances excluded the conformation in which the extended

dx.doi.org/10.1021/ac302860m | Anal. Chem. 2013, 85, 1597−1604

loop points away from the globular core as seen in the crystal structure of NKR-P1A.[10] However, the cross-link between loop lysine residues Lys166 and Lys179 was not observed although these two residues are close to each other in the model. Since the absence of cross-links can point on some structural details,[51] we propose the following explanation. The absence of cross-link between Lys166 and Lys179 on one side and existing cross-links between Lys166 and Lys196 or Lys179 and Lys196 on the other side may be explained by a hydrogen bond network between positively charged side chains of Lys166 or Lys179 and negatively charged aspartic acid Asp176 (Figure 3c). Hydrogen bonding between Lys166 and Asp176 will leave Lys179 reactive toward Lys196 and vice versa, and bonding between Lys179 and Asp176 will allow formation of cross-link between Lys166 and Lys196.[26] In addition, such conformation would disfavor cross-linking between Lys166 and Lys176. Indeed, as no cross-links between these two residues were ever detected, this interaction should be stable in time and in solution. Furthermore, Lys166 is part of a β-sheet that plays a major role in the C-type lectin-like fold and thus constrains its fluctuation. Finally, all experimentally determined distances are correctly represented in the structural model and are consistent with the ability of DSS and DSG to cross-link primary amines on residues. The extended loop hereby seems to be packed onto the protein surface partially stabilizing the second β-sheet, which is probably the most populated conformation in solution.

Native Mass Spectrometry and Ion Mobility Measurement. In order to verify the fold of NKR-P1C, native mass spectrometry with ion mobility measurements were performed. Under native conditions, the protein sprays as a monomer (Figure 4a) with main charge states ranging from 6+ to 8+. The measurement of the drift time under increasing acceleration voltages for these charge states (Figure 4b) enabled us to calculate a collisional cross section (CCS) for each one of these charge states. The values for the charge states 6+, 7+, and 8+ were 1444 ± 10 Å^2, 1495 ± 16 Å^2, and 1491 ± 16 Å^2, respectively. The monomodal and symmetrical distributions of the arrival time distributions (ATD) (Figure 4c) strongly suggest that the protein is present in a unique, very homogeneous conformation. The value obtained for the lowest charge state (1444 Å^2), which is usually the least activated in the gas phase,[52] was very close to the theoretical CCS calculated for our model (1411 Å^2, Figure 4c). As a comparison, the more extended X-ray structure obtained for the homologue NKR-P1A (pdb-code 3M9Z) shows a much higher theoretical CCS of 1618 Å^2. The difference between these two CCS (~15%) is well above the usual experimental error (±6%).[40]

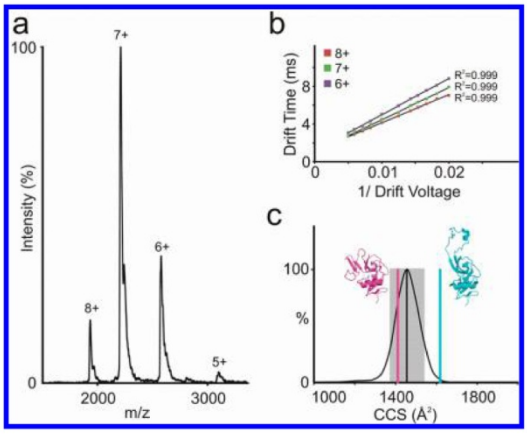


Figure 4. Ion mobility-mass spectrometry: (a) mass spectra of NKR-P1C under native conditions (200 mM ammonium acetate pH 7.4). (b) Plot of the drift times obtained for three different charge states and increasing drift voltages. (c) Arrival time distribution (ATD) obtained for the 6+ charge state (the black line represents the average ATD and the shaded area the ±6% experimental error). The pink and blue lines correspond to the theoretical CCS calculated for the model of NKR-P1C and its homologue NKR-P1A (pdb-code 3M9Z), respectively.

In this study, we combined molecular modeling together with chemical cross-linking and ion-mobility mass spectrometry to describe the structure of the extracellular part of an NK cell receptor NKR-P1C. Despite its high sequence similarity with NKR-P1A, all attempts to solve the high-resolution structure of NKR-P1C using X-ray crystallography failed. The crystals either did not grow or were of poor quality. Therefore we decided to use a combination of low-resolution MS-based methods with homology modeling to obtain the structure of this important lymphocyte receptor.

First, we obtained distance constraints using disulfide mapping and chemical cross-linking of NKR-P1C using amine-reactive cross-linkers DSS and DSG. In disulfide mapping, we first measured the mass of the intact protein by high-resolution ESI-FTICR MS. The experimental mass fitted very well to the theoretical mass with all six cysteine residues in oxidized form, confirming that the protein contains three disulfide bridges. Detailed LC−MS analysis of the disulfide-linked peptides generated from the Asp-N and tryptic digest revealed the unambiguous disulfide pattern. The identified disulfide arrangement in NKR-P1C is in agreement with the highly conserved disulfide pattern of C-type lectin-like domain[53] and with putative disulfide linkages reported previously[43] (Figure 1 and Table 1). Chemical cross-linking using nondeuterated and deuterated cross-linking agents[54] provided seven distances for DSS and four nonredundant constraints for DSG. Three DSS cross-links were cross-confirmed from the DSG experiment, and for the four remaining DSS cross-links the equivalents of DSG cross-links were not identified. This observation is consistent with the fact that cross-linkers with longer spacer arms are more likely to form cross-links compared to shorter ones.[55] However, it is unrealistic to combine data from cross-linkers with different spacer arms in order to determine lower boundaries for space distances in a quantitative way.[56]

Restraint-based computational modeling was then used to generate a model that represents the experimentally determined constraints with a minimum of violations. Molecular dynamics was used to refine the model and to describe the most populated protein conformers in solution. In addition to the positional constraints obtained from the disulfide mapping and from cross-linking experiments, the model needs to preserve the overall C-type lectin-like fold in these simulations, as the protein core is strongly conserved, and the template and modeled structure share a sequence identity of 88%. However, as crystal structures are rigid contrary to protein dissolved in solution, we allowed the side chains in the core to be more flexible and adapt to the given experimentally determined constraints. Specific attention was paid to the extended loop region (P161−G187) proposed to be involved in protein−ligand interactions and ligand specificity. The only crystal structure published to date for the entire NKR-P1 family,

mouse NKR-P1A, shows this extended loop pointing away from the protein core, in a conformation in which the loop would be fully exposed to the solvent. Such a conformation could be clearly excluded from the cross-links of the protein in solution. This was further supported by IM-MS measurements corresponding to the compact form of the molecule based on the experimentally derived collisional cross section. Therefore, in the most populated conformation in solution, NKR-P1C most likely adopts the conformation similar to the solution structure of NKR-P1A. Our model enables us to describe this conformation on an atomic scale, showing its packing onto the protein surface while partially stabilizing the second β-sheet. This demonstrates not only the principal high flexibility of this region, with conformations of the extended loop ranging from full solvent exposure to a close attachment to the protein core. As this loop is proposed to play a key role in ligand specificity and varies in primary sequence in the NKR-P1 family members, we might consider its concrete dynamical behavior as being not a collective property of the whole family but rather a specific property of NKR-P1C. This loop might thus contribute to ligand-binding specificity and consequently to the differentiation between activating and inhibitory receptors.

In conclusion, this is the first report describing combined use of chemical cross-linking, homology modeling, and ion mobility mass spectrometry for the purpose of structure elucidation of a protein with no known high-resolution structure. We show that the structure of the C-type lectin-like domain of NKR-P1C receptor is in solution similar to the homologous protein NKR-P1A and that its loop region is attached to the protein core. This also nicely demonstrates how ion mobility mass spectrometry can be used in validation of low-resolution structures in cases where different possible conformations differ in their cross sections.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: pnovak@biomed.cas.cz. Phone: +420 2 4106 2156.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by Charles University Grant Agency (Grant 403211/2011), Charles University (Project UNCE_204025/2012), the Grant Agency of the Czech Republic (Grants P207/10/1040, P303/09/0477, and P207/10/1934), the Ministry of Education, Youth and Sports of the Czech Republic (Grant MSM6007665808), University of South Bohemia (Grant GAJU 170/2010/P), and the Academy of Sciences of the Czech Republic (Grant AVOZ60870520).

■ REFERENCES

(1) Vivier, E.; Raulet, D. H.; Moretta, A.; Caligiuri, M. A.; Zitvogel, L.; Lanier, L. L.; Yokoyama, W. M.; Ugolini, S. *Science* **2011**, *331*, 44−49.

(2) Glimcher, L.; Shen, F. W.; Cantor, H. *J. Exp. Med.* **1977**, *145*, 1−9.

(3) Koo, G. C.; Peppard, J. R. *Hybridoma* **1984**, *3*, 301−303.

(4) Karlhofer, F. M.; Yokoyama, W. M. *J. Immunol.* **1991**, *146*, 3662−3673.

(5) Reichlin, A.; Yokoyama, W. M. *Immunol. Cell Biol.* **1998**, *76*, 143−152.

(6) Ljutic, B.; Carlyle, J. R.; Filipp, D.; Nakagawa, R.; Julius, M.; Zuniga-Pflucker, J. C. *J. Immunol.* **2005**, *174*, 4789−4796.

(7) Ryan, J. C.; Turck, J.; Niemi, E. C.; Yokoyama, W. M.; Seaman, W. E. *J. Immunol.* **1992**, *149*, 1631−1635.

(8) Giorda, R.; Trucco, M. *J. Immunol.* **1991**, *147*, 1701−1708.

(9) Sovova, Z.; Kopecky, V., Jr.; Pazderka, T.; Hofbauerova, K.; Rozbesky, D.; Vanek, O.; Bezouska, K.; Ettrich, R. *J. Mol. Model.* **2011**, *17*, 1353−1370.

(10) Kolenko, P.; Rozbesky, D.; Vanek, O.; Kopecky, V., Jr.; Hofbauerova, K.; Novak, P.; Pompach, P.; Hasek, J.; Skalova, T.; Bezouska, K.; Dohnalek, J. *J. Struct. Biol.* **2011**, *175*, 434−441.

(11) Rozbesky, D.; Man, P.; Kavan, D.; Chmelik, J.; Cerny, J.; Bezouska, K.; Novak, P. *Anal. Chem.* **2012**, *84*, 867−870.

(12) Kaltashov, I. A.; Bobst, C. E.; Abzalimov, R. R.; Berkowitz, S. A.; Houde, D. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 323−337.

(13) Serpa, J. J.; Parker, C. E.; Petrotchenko, E. V.; Han, J.; Pan, J.; Borchers, C. H. *Eur. J. Mass Spectrom. (Chichester, Eng.)* **2012**, *18*, 251−267.

(14) Rezabkova, L.; Man, P.; Novak, P.; Herman, P.; Vecer, J.; Obsilova, V.; Obsil, T. *J. Biol. Chem.* **2011**, *286*, 43527−43536.

(15) Lasker, K.; Phillips, J. L.; Russel, D.; Velazquez-Muriel, J.; Schneidman-Duhovny, D.; Tjioe, E.; Webb, B.; Schlessinger, A.; Sali, A. *Mol. Cell. Proteomics* **2010**, *9*, 1689−1702.

(16) Stengel, F.; Aebersold, R.; Robinson, C. V. *Mol. Cell. Proteomics* **2012**, *11*, R111 014027.

(17) Young, M. M.; Tang, N.; Hempel, J. C.; Oshiro, C. M.; Taylor, E. W.; Kuntz, I. D.; Gibson, B. W.; Dollinger, G. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5802−5806.

(18) Jacobsen, R. B.; Sale, K. L.; Ayson, M. J.; Novak, P.; Hong, J.; Lane, P.; Wood, N. L.; Kruppa, G. H.; Young, M. M.; Schoeniger, J. S. *Protein Sci.* **2006**, *15*, 1303−1317.

(19) Ahrman, E.; Lambert, W.; Aquilina, J. A.; Robinson, C. V.; Emanuelsson, C. S. *Protein Sci.* **2007**, *16*, 1464−1478.

(20) Schulz, D. M.; Ihling, C.; Clore, G. M.; Sinz, A. *Biochemistry* **2004**, *43*, 4703−4715.

(21) Maiolica, A.; Cittaro, D.; Borsotti, D.; Sennels, L.; Ciferri, C.; Tarricone, C.; Musacchio, A.; Rappsilber, J. *Mol. Cell. Proteomics* **2007**, *6*, 2200−2211.

(22) Chen, Z. A.; Jawhari, A.; Fischer, L.; Buchen, C.; Tahir, S.; Kamenski, T.; Rasmussen, M.; Lariviere, L.; Bukowski-Wills, J. C.; Nilges, M.; Cramer, P.; Rappsilber, J. *EMBO J.* **2010**, *29*, 717−726.

(23) Rappsilber, J. *J. Struct. Biol.* **2011**, *173*, 530−540.

(24) Novak, P.; Haskins, W. E.; Ayson, M. J.; Jacobsen, R. B.; Schoeniger, J. S.; Leavell, M. D.; Young, M. M.; Kruppa, G. H. *Anal. Chem.* **2005**, *77*, 5101−5106.

(25) Bruce, J. E. *Proteomics* **2012**, *12*, 1565−1575.

(26) Novak, P.; Kruppa, G. H.; Young, M. M.; Schoeniger, J. *J. Mass Spectrom.* **2004**, *39*, 322−328.

(27) Novak, P.; Kruppa, G. H. *Eur. J. Mass Spectrom. (Chichester, Eng.)* **2008**, *14*, 355−365.

(28) Matsumura, M.; Signor, G.; Matthews, B. W. *Nature* **1989**, *342*, 291−293.

(29) Wedemeyer, W. J.; Welker, E.; Narayan, M.; Scheraga, H. A. *Biochemistry* **2000**, *39*, 4207−4216.

(30) Pitt, J. J.; Da Silva, E.; Gorman, J. J. *J. Biol. Chem.* **2000**, *275*, 6469−6478.

(31) Jensen, P. K.; Harrata, A. K.; Lee, C. S. *Anal. Chem.* **1998**, *70*, 2044−2049.

(32) Smith, D. L.; Zhou, Z. R. *Methods Enzymol* **1990**, *193*, 374−389.

(33) Wallis, T. P.; Pitt, J. J.; Gorman, J. *J. Protein Sci.* **2001**, *10*, 2251−2271.

(34) Pompach, P.; Man, P.; Kavan, D.; Hofbauerova, K.; Kumar, V.; Bezouska, K.; Havlicek, V.; Novak, P. *J. Mass Spectrom.* **2009**, *44*, 1571−1578.

(35) McLafferty, F. W.; Fridriksson, E. K.; Horn, D. M.; Lewis, M. A.; Zubarev, R. A. *Science* **1999**, *284*, 1289−1290.

(36) Smith, R. D.; Loo, J. A.; Loo, R. R. O.; Busman, M.; Udseth, H. R. *Mass Spectrom. Rev.* **1991**, *10*, 359−451.
(37) Ruotolo, B. T.; Giles, K.; Campuzano, I.; Sandercock, A. M.; Bateman, R. H.; Robinson, C. V. *Science* **2005**, *310*, 1658−1661.
(38) Lane, L. A.; Fernandez-Tornero, C.; Zhou, M.; Morgner, N.; Ptchelkine, D.; Steuerwald, U.; Politis, A.; Lindner, D.; Gvozdenovic, J.; Gavin, A. C.; Muller, C. W.; Robinson, C. V. *Structure* **2011**, *19*, 90−100.
(39) Baldwin, A. J.; Lioe, H.; Hilton, G. R.; Baker, L. A.; Rubinstein, J. L.; Kay, L. E.; Benesch, J. L. *Structure* **2011**, *19*, 1855−1863.
(40) Hall, Z.; Politis, A.; Bush, M. F.; Smith, L. J.; Robinson, C. V. *J. Am. Chem. Soc.* **2012**, *134*, 3429−3438.
(41) Hilton, G. R.; Thalassinos, K.; Grabenauer, M.; Sanghera, N.; Slade, S. E.; Wyttenbach, T.; Robinson, P. J.; Pinheiro, T. J.; Bowers, M. T.; Scrivens, J. H. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 845−854.
(42) Bagal, D.; Valliere-Douglass, J. F.; Balland, A.; Schnier, P. D. *Anal. Chem.* **2010**, *82*, 6751−6755.
(43) Rozbesky, D.; Kavan, D.; Chmelik, J.; Novak, P.; Vanek, O.; Bezouska, K. *Protein Expression Purif.* **2011**, *77*, 178−184.
(44) Sali, A.; Blundell, T. L. *J. Mol. Biol.* **1993**, *234*, 779−815.
(45) Watson, A. A.; Lebedev, A. A.; Hall, B. A.; Fenton-May, A. E.; Vagin, A. A.; Dejnirattisai, W.; Felce, J.; Mongkolsapaya, J.; Palma, A. S.; Liu, Y.; Feizi, T.; Screaton, G. R.; Murshudov, G. N.; O'Callaghan, C. A. *J. Biol. Chem.* **2011**, *286*, 24208−24218.
(46) Guex, N.; Peitsch, M. C. *Electrophoresis* **1997**, *18*, 2714−2723.
(47) Krieger, E.; Koraimann, G.; Vriend, G. *Proteins* **2002**, *47*, 393−402.
(48) Hernandez, H.; Robinson, C. V. *Nat. Protoc.* **2007**, *2*, 715−726.
(49) Bush, M. F.; Hall, Z.; Giles, K.; Hoyes, J.; Robinson, C. V.; Ruotolo, B. T. *Anal. Chem.* **2010**, *82*, 9557−9565.
(50) Benesch, J. L. P.; Ruotolo, B. T. *Curr. Opin. Struct. Biol.* **2011**, *21*, 641−649.
(51) Novak, P.; Young, M. M.; Schoeniger, J. S.; Kruppa, G. H. *Eur. J. Mass Spectrom. (Chichester, Eng.)* **2003**, *9*, 623−631.
(52) Clemmer, D. E.; Jarrold, M. F. *J. Mass Spectrom.* **1997**, *32*, 577−592.
(53) Zelensky, A. N.; Gready, J. E. *FEBS J.* **2005**, *272*, 6179−6217.
(54) Muller, D. R.; Schindler, P.; Towbin, H.; Wirth, U.; Voshol, H.; Hoving, S.; Steinmetz, M. O. *Anal. Chem.* **2001**, *73*, 1927−1934.
(55) Leitner, A.; Walzthoeni, T.; Kahraman, A.; Herzog, F.; Rinner, O.; Beck, M.; Aebersold, R. *Mol. Cell. Proteomics* **2010**, *9*, 1634−1649.
(56) Back, J. W.; de Jong, L.; Muijsers, A. O.; de Koster, C. G. *J. Mol. Biol.* **2003**, *331*, 303−313.