THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Structural analyses of immune cell receptor signalling and activation

MARIA SALINE

GÖTEBORGS UNIVERSITET

University of Gothenburg
Department of Chemistry and Molecular Biology
Göteborg, Sweden, 2013

Structural analyses of immune cell receptor signalling and activation.

Maria Saline

Cover: Artistic interpretation of the whole T-cell receptor complex, superantigen and MHC in front of the NMR magnet. Based on *Paper I* and *IV* [1].

By: Isabelle Fellbom

ISBN 978-91-628-8683-7

Available online at http://hdl.handle.net/2077/32556

University of Gothenburg
Department of Chemistry and Molecular Biology
SE-405 30 Gothenburg
Sweden
Telephone: +46 (0)31-786 0000

ABSTRACT

Structural biology is a scientific field where the aim is to observe macromolecules on a atomic level to understand their functions. Often these macromolecules are proteins that make, almost, everything happen within the cells; from hormones and enzymes to building blocks and controlling gene expression. With structural biology tools, scientists can visualize structures, interactions and mobility within the proteins to get an insight in the chain of events in a cell.

In this thesis, the function of signalling has been in focus. Both signalling within the immune system through a membrane bound receptor that finds an invading pathogen and signals into the cells “alert, we have an invader” where the immune system reacts. How the signalling passes from the outside of the cell to the inside is still not revealed. One part of this thesis investigates the outside of an immune cell and the other, the inside, past the membrane.

Proteins are chains of amino acids that are predestined to either fold into a stable structure or stay loose and flexible. Superantigens are very stable proteins; they are toxins, made to last and conquer. The opposite are the intracellular flexible domains of the immune receptors which belong to a class of proteins, so-called intrinsically disordered proteins, IDPs, which are less investigated but omnipresent. In this thesis some flexible domains of the immune receptors have efficiently been produced in a cell free protein synthesis and examined by NMR, using a new setup of acquisition and analysis. All domains are lacking secondary structure and a well-defined three-dimensional structure. The proteins investigated more in depth within the work of this thesis, show tendency for α-helical regions, most likely of functional significance.

Viruses are evolved to use its host and get a free ride. Here we explore the interaction of one SIV (orthologous to HIV) protein with one of the intracellular flexible domains of the T-cell receptor, which leads to down regulation of the receptor resulting in immune deficiency. This interaction is unique in that no changes in the very sensitive NMR spectra are seen; yet other techniques indicate specific interaction.

As the SIV protein is abusing the immune system, superantigens hijack the immune system by crosslinking the T-cell receptor to an antigen-presenting cell displaying pieces of an invading pathogen on it´s surface, and by this start an extreme immune response, sometimes lethal. This superantigen can circumvent the intricate, specific and effective immune system and they are up to date thought to interact with the β-chain of the T-cell receptor. We show in this thesis by structural biology techniques such as x-ray and NMR that a superantigen interacts with the α-chain of the TCR. This is the first structure structurally determined ternary complex of an antigen-MHC-superantigen-TCR, a paradigm shift in superantigen biology.

LIST OF ORIGINAL PUBLICATIONS

The thesis is based on the following publications, which will be referred to in the text by their Roman numerals:

Paper I Sigalov AB, Kim WM, Saline M, Stern LJ. (2008) The intrinsically disordered cytoplasmic domain of the T cell receptor zeta chain binds to the nef protein of simian immunodeficiency virus without a disorder-to-order transition. *Biochemistry*. Dec 9;47(49):12942-4.

Paper II Saline M, Orekhov V, Lindkvist-Petersson K, Karlsson BG. (2010) Backbone resonance assignment of Staphylococcal Enterotoxin H. *Biomolecular NMR Assignment* Apr;4(1):1-4. Epub 2009 Nov 4.

Paper III Saline M, Rödström KE, Fischer G, Orekhov VY, Karlsson BG, Lindkvist-Petersson K. (2010) The structure of superantigen complexed with TCR and MHC reveals novel insights into superantigenic T cell activation. *Nat Communications* Nov;1(8):119.

Paper IV Isaksson L, Mayzel M, Saline M, Pedersen A, Rosenlöw J, Brutscher B, Karlsson BG, Orekhov VY. (2013) Highly efficient NMR assignment of intrinsically disordered proteins: application to B- and T cell receptor domains. *Accepted for publication in PLoS ONE.*

Contribution report:

I: I performed an analysed the NMR experiments as well as SDS-PAGE to confirm presence of the proteins in the NMR tube.

II: I expressed, purified and ran NMR experiments on SEH. I performed full backbone assignment as well and took part in writing the paper.

III: I expressed and purified all proteins, ran and analysed all NMR experiments. I analysed X-ray results and took part in writing the paper.

IV: I cloned, expressed and purified the cytoplasmic domains. I ran NMR experiments, analysed data and took part in writing the paper.

ABREVIATIONS

BCR	B lympocyte/B-cell Receptor
BMRB	BioMolecularResonanceBank
CDR1-3	Complementary determining regions of TCR variable chain
CDx	Cluster of Differentiation x
FR	Framework region of TCR variable chain
HSQC	Heteronuclear single quantum coherence spectroscopy
HV	Hyper variable loop of TCR variable chain
IDP	Intrinsically Disordered Protein
K_d	Equilibrium constant, in unit M
kDa	kilo Dalton, g/mol, unit for size of proteins
M (μM)	Molar, mol/dm^3, unit for concentration
MHC	Major Histocompatibiliyt Complex
NMR	Nuclear Magnetic Resonance
PDB	Protein data bank
PTK	Protein tyrosine kinase
SAg	Superantigen
SEH	Staphylococcal Enterotoxin H
TCR	T lymphocyte/T-cell Receptor
TSST	Toxic Shock Syndrome Toxin, a SAg
Vα	Variable chain α of TCR
Vβ	Variable chain β of TCR

TABLE OF CONTENTS

BACKGROUND **1**

IMMUNE SYSTEM 1

General *1*

B-lymphocyte *2*

T-Cell Receptor, B-Cell Receptor *4*

SUPERANTIGENS 7

INTRINSICALLY DISORDERED PROTEINS 10

Protein complexes of intrinsically disordered proteins *11*

In the cell *12*

Disease and IDP *12*

Intracellular domains of TCR *15*

Intracellular domains of BCR *16*

SIGNALLING **17**

SIGNALLING IN THE TCR AND BCR 17

In the membrane/segregation: *18*

Aggregation: *20*

Conformational change: *22*

The role of the intracellular disordered domains: *22*

SIGNALLING IN CELLS INFECTED BY HIV/SIV 24

SIGNALLING INDUCED BY SUPERANTIGENS 25

Superantigen binding to MHC *25*

Superantigens binding to TCR *25*

Detailed analysis of TCR-SAg interactions *27*

The ternary complexes MHC-SAg-TCR *27*

Affinity maturated SAg *28*

Activation by superantigens *28*

STRUCTURAL BIOLOGY **30**

PROTEIN PRODUCTION 31

In vivo protein expression *31*

Cell-free protein synthesis (CFPS) *32*

BIOPHYSICAL TECHNIQUES 35
X-Ray Crystallography *35*
NMR Spectroscopy *37*
Interactions investigated by NMR *42*
OTHER BIOPHYSICAL TECHNIQUES FOR INVESTIGATIONS OF INTRINSICALLY DISORDERED PROTEINS 44
Computational methods *44*
Biochemical/biophysical methods *45*
SUMMARY OF PUBLICATIONS: **48**
PAPER I; "THE INTRINSICALLY DISORDERED CYTOPLASMIC DOMAIN OF THE T CELL RECEPTOR ZETA CHAIN BINDS TO THE NEF PROTEIN OF SIMIAN IMMUNODEFICIENCY VIRUS WITHOUT A DISORDER-TO-ORDER TRANSITION." 48
PAPER II; "BACKBONE RESONANCE ASSIGNMENT OF STAPHYLOCOCCAL ENTEROTOXIN H." 49
PAPER III; "THE STRUCTURE OF SUPERANTIGEN COMPLEXED WITH TCR AND MHC REVEALS NOVEL INSIGHTS INTO SUPERANTIGENIC T CELL ACTIVATION." 50
PAPER IV; "HIGHLY EFFICIENT NMR ASSIGNMENT OF INTRINSICALLY DISORDERED PROTEINS: APPLICATION TO B- AND T CELL RECEPTOR DOMAINS" 51
ACKNOWLEDGEMENT **52**
REFERENCES **55**

"Without some disorder, nothing can be alive"

Professor RJP Williams, University of Oxford

Well-renowned NMR spectroscopist, member of numerous science academies

BACKGROUND

Immune system

General

The immune system protects us from all kind of invading pathogens; bacteria, viruses, microorganisms and it is very effective and complex. In principle, it can be divided into innate and adaptive immunity where the first is nonspecific such as anatomic barriers; the skin and mucus, physiological; low pH in the stomach, endocytocis/phagocytosis; digestion of microorganisms and inflammation; increase of the permeability of the capillaries at the site of infection and recruitment of phagocytes. The second, adaptive part of the immune system is very specific and responsive to full or digested pieces of microorganisms, so called antigens, with specificity and diversity [2], see *figure 1*. The adaptive immune system is capable of generating enormous diversity through gene rearrangement of the multi chain immune recognition receptors (MIRR) in the maturation process of the cells and to recognize billions of different antigens [3]. Once an immune cell recognizes and responds to an antigen, a massive amount of that particular immune cell is produced by so-called clonal expansion and it also starts an immune response, which includes production of cytokines, release of antibodies, cell killing (cytotoxicity) and the creation of immunologic memory [4, 5]. The two major groups of cells active in the adaptive immune system are the lymphocytes that mature in the bone marrow (B-lymphocytes) or in the thymus (T-lymphocytes) where they are selected for specific affinity for only non-self antigens, i.e. foreign molecules. If the maturation process fails we will have reaction against self-antigens, molecules within our bodies and this may lead to autoimmune diseases [6]. The opposite might happen; the immune system is not strong enough and fails to observe and fight a virus, bacterium or microorganism, so called immunodeficiency, such as HIV/AIDS [7]. Another case, when it goes wrong is the case of engagement by superantigens [8] where the immune system overreacts and we get sick (*Papers II, III*).

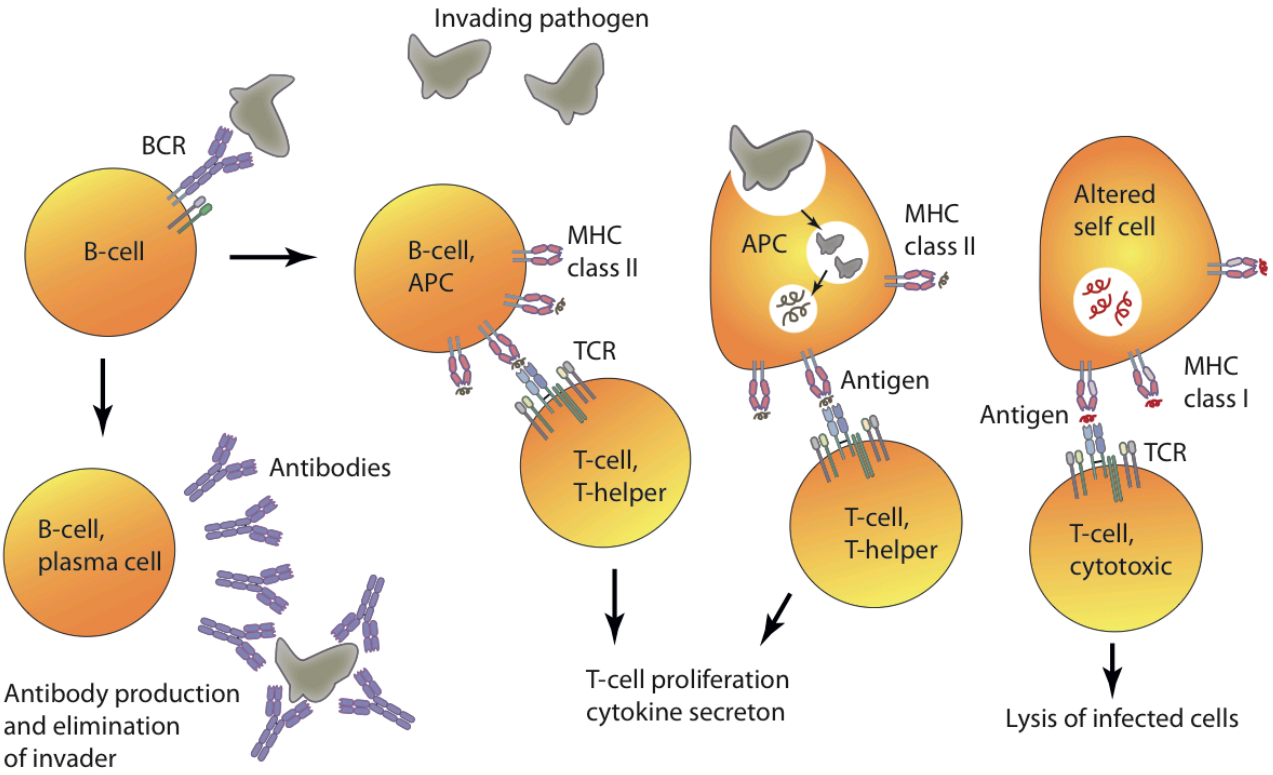


Figure 1. The adaptive immune system can yet be divided into humoral and cell mediated branches, where the first is the interaction of antigens leading to activated B-cells that produce soluble antibodies (immunoglobulines, Ig) which bind to antigens, neutralizes them and act as effectors in the humoral system. The cell mediated branch of the adaptive immune system targets both self-cells that have been infected with viruses and express a molecule called Major Histocompatibility Complex, MHC, class I and a group of immune cells called Antigen-Presenting Cells, APC, which are e.g. macrophages or B-lymphocytes that present antigens on their surface. APCs that recognize invaders, usually bacteria, take them into the cell through endocytosis and process them to smaller pieces, antigens and presents them using MHC class II [2]. Both B-cells and T-cells recognize the antigen presented by altered self-cells and APC. Co-receptors are utilized for better and correct binding to the lymphocytes and at the binding to the proper B-cell receptor (BCR) or T-cell receptor (TCR) the lymphocyte starts to proliferate and act on the invader.

B-lymphocyte

B-cells are both APC and antigen binding cells at the same time. The surface bound antigen binding receptor develops from pro-BCR to pre-BCR and eventually the mature B-cell receptor, which is ready to recognize antigens; antigens soluble in the solution, haptens with carrier proteins or antigens presented by APC [9]. BCR recognizes antigens of different sizes and variety; they can be lipids or carbohydrates as well as proteins. Upon

antigen binding the BCR can either internalize, processes the antigen bound and presents it for TCR to find and respond to or turn into a so-called plasma cell. The plasma cell produce soluble immunoglobulins (Ig), also called antibodies, identical to the extracellular part of the receptor initially recognizing the antigen, and these Ig have a strong affinity for that special pattern of the invader and are important molecules in the elimination process as well as in immunological memory: immunity [2].

T-lymphocyte

There are two major groups of T-cells, cytotoxic and helper cells. They both require the antigen presented by a MHC molecule for activation, and the ternary complex antigen-MHC-TCR is created. In the case of altered self-cells displaying MHC class I, who binds to cytotoxic T-lymphocytes (with the co-receptor $CD8^+$), the altered self-cell is eliminated by cell-killing or cytotoxic activity. For bacterial infections, the MHC class II are recognised by T-helper cells (with the co-receptor $CD4^+$) that start to secrete cytokines and interleukins, molecules that enhances the rest of the immune system, e.g. by activating B-cells for antibody production [2, 10].

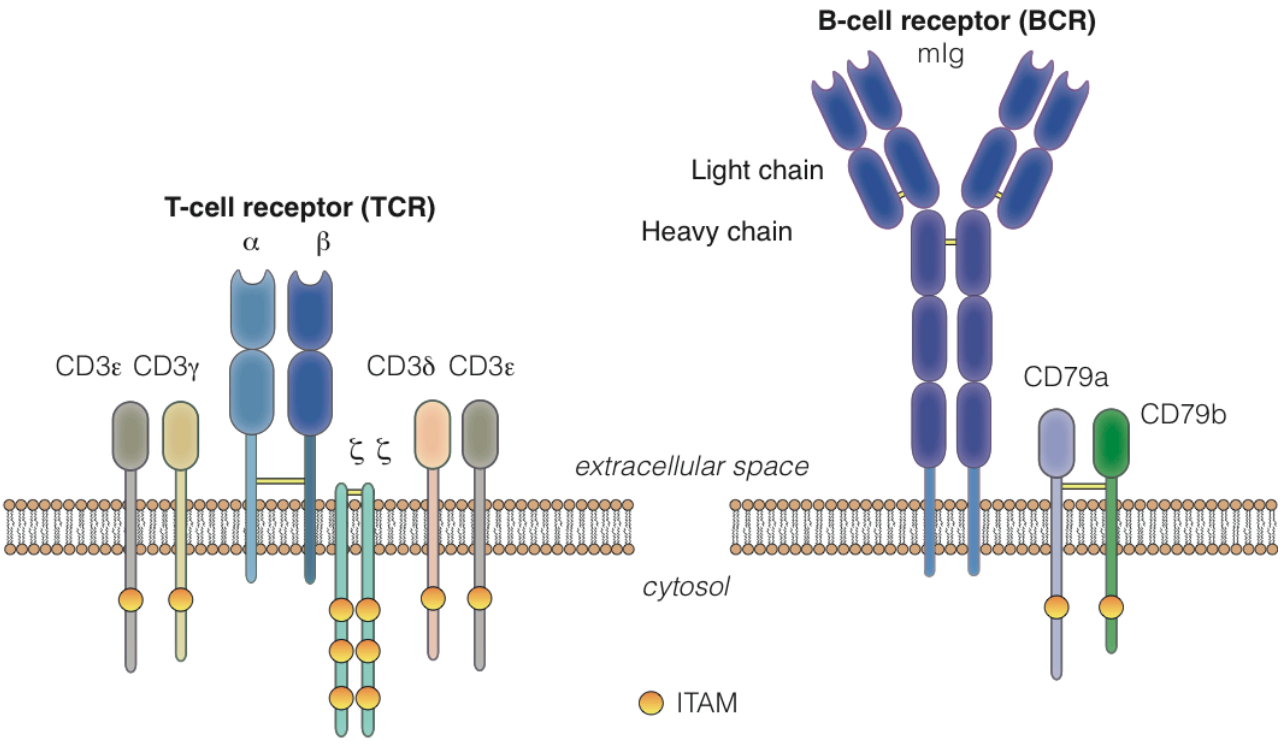


Figure 2. Cartoon of The T-cell receptor and B-cell receptor complexes. The complete T-cell receptor is a complex of TCRαβ that associates with the CD3 complex. CD3 itself is made up of five invariant polypeptide chains that makes up three dimers; εδ, εγ and ζζ. All proteins in the TCR complex have to be

expressed for the receptor to be displayed on the surface [11]. These CD3 chains all have an intrinsically disordered intracellular chain with the common ITAM sequence and εδ and εγ have an extracellular globular domain, similar to the immunoglobulin fold (PDB: 1XMW, 1JBJ) [12, 13] and they have a transmembrane α-helix connecting them through the membrane. The CD3 complex is associated to TCRαβ both on extracellular side [14] as well as via a complex network of electrostatic interactions of the α-helices in the membrane [15]. Acidic residues in the CD3 dimers and basic amino acids on the TCRαβ hold the complex together [16]. Exactly how these domains are arranged in the membrane is not fully known; most recently the thought is that CD3ε of both dimers are facing each other and all three CD3 dimers on one side of the TCRαβ enabling dimerization of the TCRs in the membrane upon antigen binding [17].

Like TCR, the BCR heavy and light chain dimer interacts with the molecules that connects to the outside signal to the inside of the celland these are the heterodimer CD79a/ CD79b. They both have an extracellular Ig-domain, an α-helix through the membrane and an intracellular intrinsically disordered domain containing ITAM. The CD79a/CD79b are bound together by a disulphide-bond on the extracellular side and are connected to mIg in the membrane. Picture adapted from *Paper IV*.

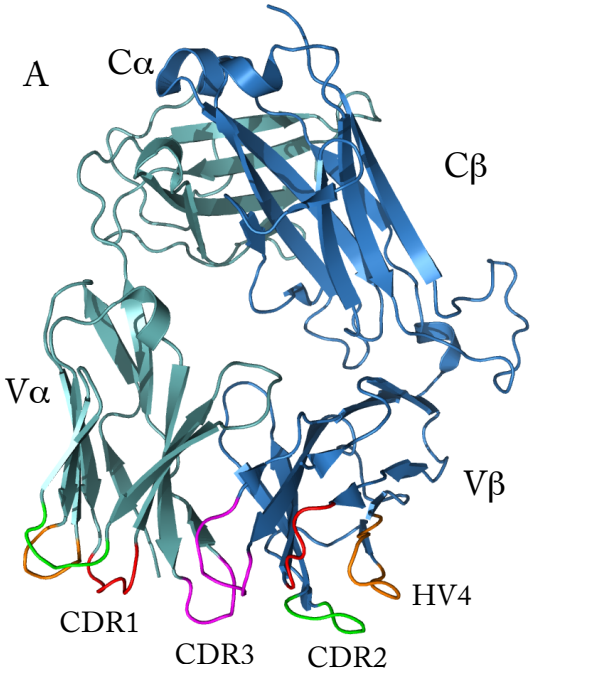
T-Cell Receptor, B-Cell Receptor

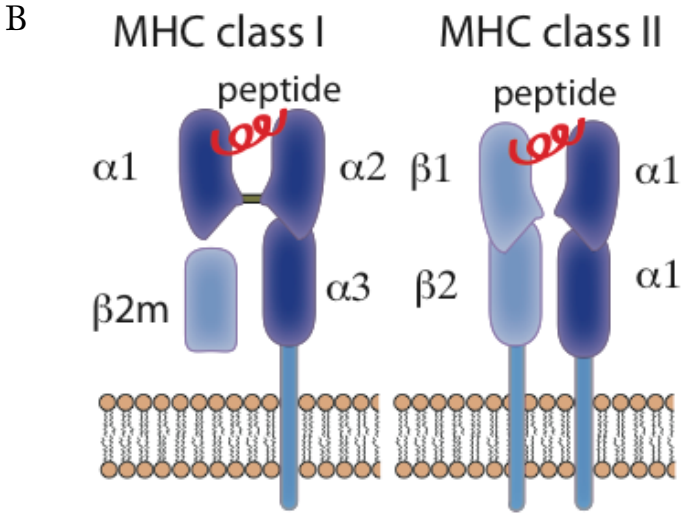
On the surface of the B- and T-cells, there are several different receptors to aid the specificity of interaction and to keep cells attached to each other during antigen presentation. Focus of this thesis has been on the receptors that recognize antigens; they are of the group multichain immune-recognition receptor, MIRR, and are called the B-cell and T-cell receptor. Several things are in common for these proteins, such as that the extracellular parts both are anchored to the membrane with a transmembrane helix that enters a few residues on the cytosolic side of the cell [18]. These few intracellular residues are not sufficient for signalling into the cell. Thus, the receptor complexes, consists of the extracellular part associated with intracellular proteins, which also are membrane bound, see *figure 2*, to make up the functional receptor complex [4, 11]. These intracellular proteins are all classified as intrinsically disordered proteins, IDPs, explained later in this thesis, and contain a specific phosphorylation motif, immune-receptor tyrosine-based activation motif, ITAM, crucial for downstream signalling in the cell. This is a conserved sequence of amino acids with two tyrosines (Y) that are targets for kinases which adds and phosphate and to the tyrosines, signalling activation [19]. The extracellular domains consist of a variable and a constant region where the variable, through gene rearrangement in the maturation process (and somatic mutation for BCR), gives rise to $10^{10}/10^{8}$ (BCR/TCR) different epitopes that can recognize as many antigens. To pass the positive selection in the maturation process, TCR has to have affinity for self-peptide-MHCs and the negative selection makes sure that

this affinity is not high. However, the binding of self-peptide-MHC is a necessary event for maintenance of the mature TCR repertoire [20, 21].

The BCR complex is a membrane bound immunoglobulin, mIg, which in the membrane associates with the heterodimer CD79a/CD79b (Igα/Igβ) [22]. Each B-lymphocyte contains up to 120.000 BCR on its surface, all with the same mIg [23]. The mIg can be of several different Ig; IgA, IgD IgE, IgG, or IgM and they only differ from their soluble counterpart by the membrane spanning helix. mIgM and mIgD are the immunoglobulins in mature B-cells, the others are present later in development, such as in plasma and memory cells. The mIg comprises 2 heavy and 2 light chains, where the light chain has one variable and one constant domain disulphide-linked to the heavy chain which has one variable and 4 constant domains on the extracellular side and a transmembrane helix and 3-28 intracellular residues depending on the mIg, see *figure 2*. BCR recognizes antigens of all sizes, shapes and in this, it differs a lot from TCR, which is restricted to peptides of 8-18 amino acids [2].

The extracellular TCR is made up of two chains, α and β, each comprising of a variable and a constant chain, a linker of about 20 amino acids forms to an α-helix anchored in the membrane. The constant and variable domains are of the Ig-fold, see *figure 3A*. The Ig fold consists of two β sheets, each built of an antiparallel β strands, surrounding a central hydrophobic core. One disulphide bond bridges the two sheets and this general fold is used in many proteins within the immune system, playing key roles, e.g. antibodies i.e. BCR. In human, we have 60 different alleles for Vβ (TRBV) and 47 for Vα (TRAV) [24]. In detail these variable chains consist of a framework region; FR that is mainly the 7 β-sheets and the loops connecting the β-strands. The loops are called complementary determining regions, CDR, 1-3, see *figure 3A*. CDR1-2 and a hyper variable loop, HV4, are different between the alleles whereas CDR3, differ within the allele making 10^8 different epitopes [25]. Upon recognition of a peptide-MHC complex, the CDR1 and CDR2 loops are mostly involved in recognition of MHC, whereas the CDR3 loops, the most variable, recognize the peptide. The peptide-MHC complex is bound to the TCR with a moderate to low affinity, 10^{-7}-10^{-4} M, co-receptors of different kinds strengthen the overall binding affinity [2, 10, 26].





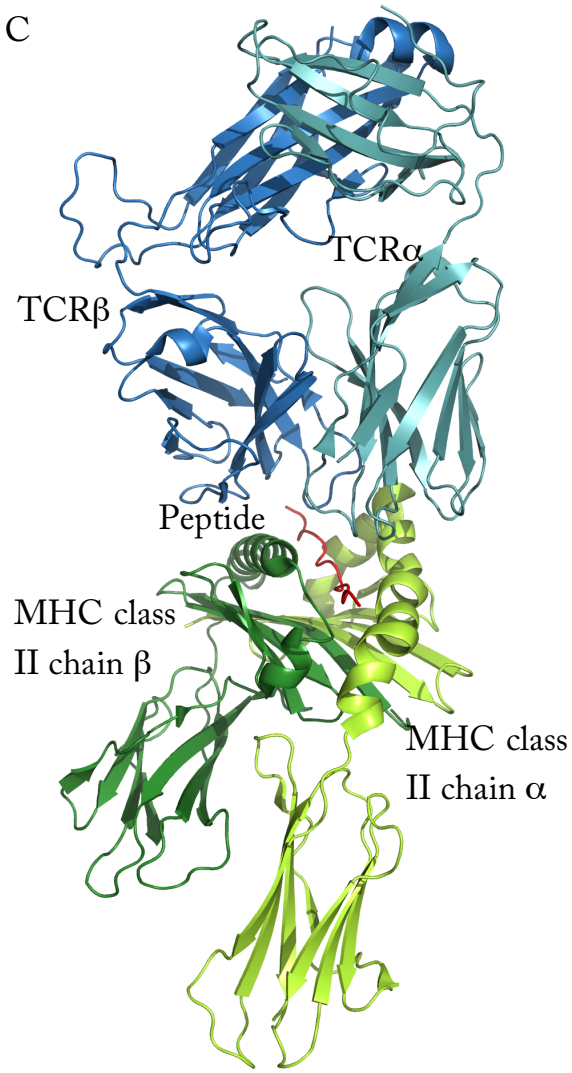


Figure 3. The variable loops in TCR, MHC cartoon and Structure of conventional peptide-MHC/TCR activation.

A; The specific loops of TCR variable domain called CDR1-3 and the HV-4, responsible for peptide binding and also the main contact points to superantigen. CDR1 – Red, CDR2 – Green, CDR3 – Purple, HV4 – orange on the TCR where the α-chain is light blue and β-chain darker blue (pdb 1OGA) [25].

B: MHC class I, II. MHC class I has a transmembrane segment and a large α-chain (45kDa) with three domains of which two make up the peptide binding groove, α1-2, and one, α3, has an immunoglobulin fold and associates to a β_2-microglobulin, β_2m, (12kDa), also of immunoglobulin fold [27]. The class II MHC has two similar chains, α (33 kDa) and β (28 kDa) which both have a immunoglobulin fold that anchors to the membrane and a domain that makes the peptide binding groove [28].

C; Conventional peptide-MHC/TCR activation. The peptide is presented by MHC class II and TCR recognizes MHC class II using CDR1 and 2. TCR interacts with the specific amino acid sequence of the antigen using the adaptable CDR3 [29] (pdb 3RDT) that show a higher mobility compared to the other regions in the variable chains [30]. Red – antigen, Green – MHC class II, Blue – TCR.

Major Histocompatibility Complex, MHC, Class I and II

The Major Histocompatibility Complexes are membrane bound proteins also called human leukocyte antigens, HLAs. They are divided in two subgroups that both present the antigen to TCR, see *figure 3B*. The two groups have different tasks and folds; class I presents peptides of 8-10 amino acids originating from the cytosol of an altered self-cell. Almost all cells in the body can produce MHC class I and they get recognized by CD8+ T-cells. Class II presents antigens of 13-18 amino acids from the exogenous pathway, presented by APC and recognized by T-cells bearing CD4+ co-receptors.

The peptide is bound with strong affinity, 10^{-5}-10^{-10} mol/L in the open groove and specific anchor amino acids in the N- and C-terminus are the most important for binding [31]. In human there are an enormous diversity in HLA genes, no individual has the same as another (identical twins excepted), and there is no gene-rearrangement as for TCR and BCR. Each individual carries up to 6 different MHC class I molecules and up to 12 MHC class II. Within one person there is not a large variety and neither acceptance for other MHCs, which give complications with transplants rejection[2]. Each MHC can bind many different antigens, and some antigens can bind several MHCs. The specificity of the immune system lies mostly in the interaction antigen-TCR selection [2].

Superantigens

In contrast to conventional antigens, superantigens, SAgs, are proteins of bacterial or viral origin that do not get processed and presented by the MHC molecules. Instead, these toxins get absorbed in the human body by the intestinal/gastric epithelium and as intact proteins they bind to both MHC class II and TCR, activating the T-cell. In a conventional T-cell response, 0.001%-0.0001% of all T-cells become activated while in the case of SAg activation up till 20% may be stimulated [32]. The reason SAg induce such great number of T-cells is that they bind outside the peptide binding groove of the MHC class II and crosslink to the TCR, leading to an extreme activation of both CD4+ and CD8+ T-cells. SAgs interact with different parts of the variable domain of TCR and some superantigens bind one other a few particular TCR-alleles. Each allele differs in the CDR3 epitopes, [33] which varies; hence, the many activated TCRs will start to proliferate and produce TCRs recognizing a wide variety of antigens but mostly a large amount of T-cells are activated.

Extremely small amounts of SAgs (pico – femtogram) in the body results in an enormous immune response with overproduction of cytokines such as interferon-γ (IFN-γ),

interleukin-2 (IL-2) and tumour necrosis factor-α (TNF-α) [34]. High levels of cytokines in the blood stream leads to diseases such as systemic inflammation and toxic shock syndrome, TSS, a potentially fatal disease, with symptoms such as vomiting, diarrhea and nausea. SAgs also causes food poisoning and are suggested to be involved in autoimmune diseases including multiple sclerosis [8] and numerous studies have also shown a role for SAgs in other diseases such as Kawasaki disease (KD), atopic dermatitis (AD), Guttate Psoriasis and chronic rhinosinusitis (CRS) [35]. A way to resolve whether a disease is due to presence of SAgs is to investigate the repertoire of TCRs in the blood stream [36]. Upregulation of one or several specific and down regulation of other alleles indicates SAg involvement.

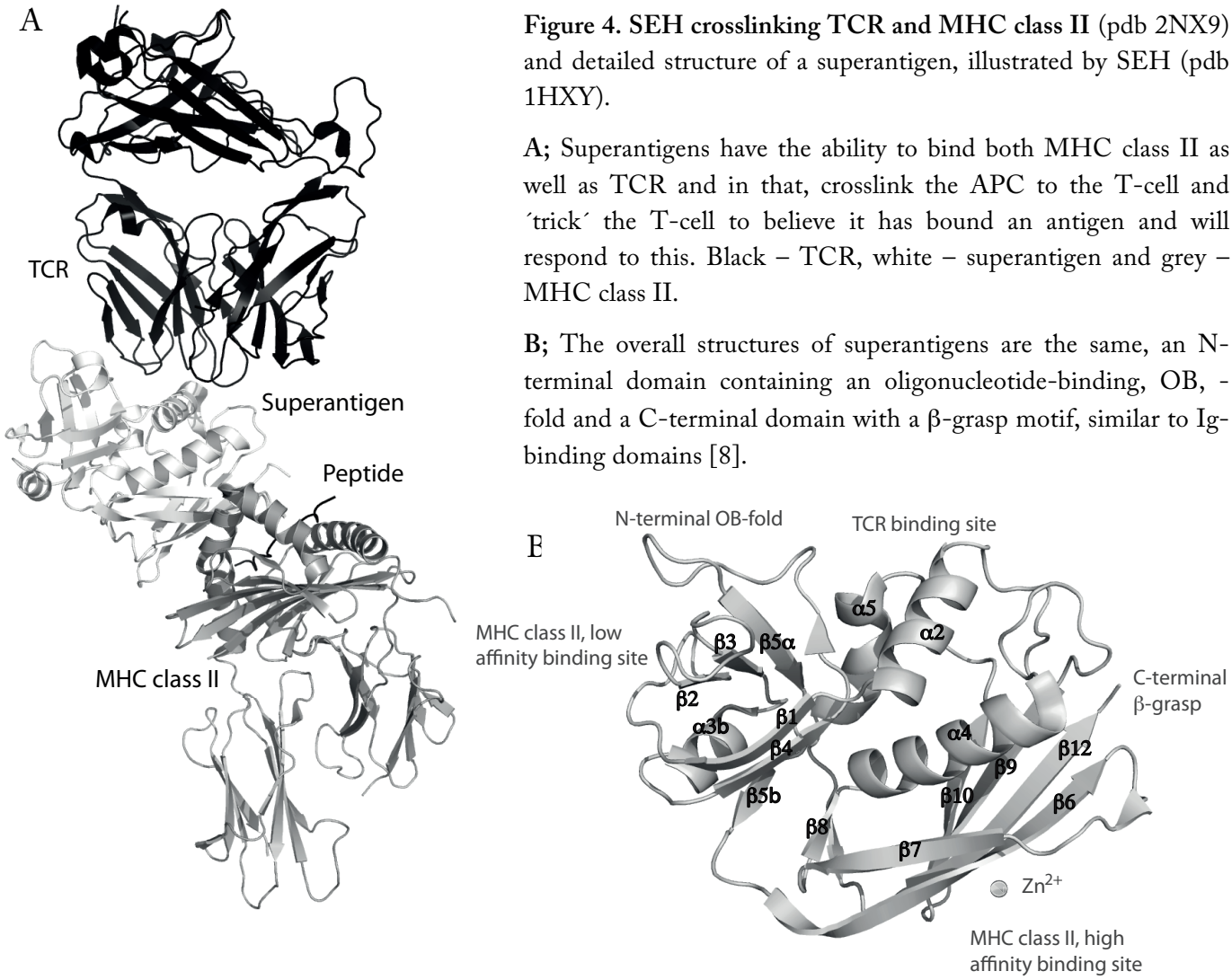


Figure 4. SEH crosslinking TCR and MHC class II (pdb 2NX9) and detailed structure of a superantigen, illustrated by SEH (pdb 1HXY).

A; Superantigens have the ability to bind both MHC class II as well as TCR and in that, crosslink the APC to the T-cell and ´trick´ the T-cell to believe it has bound an antigen and will respond to this. Black – TCR, white – superantigen and grey – MHC class II.

B; The overall structures of superantigens are the same, an N-terminal domain containing an oligonucleotide-binding, OB, -fold and a C-terminal domain with a β-grasp motif, similar to Ig-binding domains [8].

Bacterial superantigens are stable proteins ranging from 21-31kDa and these proteins are mainly produced by *Staphylococcus aureus* (SE*s*) and *Streptococcus pyogenes* (Spe*s*) and in addition *Streptococcus equi* and *Streptococcus dysgalactiae disgalactiae* both produce superantigen-like proteins (SE*lx)* lacking the ability to bind TCR and MHC but with the superantigen fold also produce superantigens (SePEs and SDM) [8, 34]. Toxins produced by *Yersinia pseudotuberculosis* (YPM) and *Mycoplasma arthritidis* (MAM) have a different fold but generally the same functions as superantigens. SSL staphylococcal superantigen-like proteins have similar folds but possess no superantigenic activity [34].

S. aureus produces several different superantigens, known to date are the staphylococcal enterotoxins SE*s* (A-E, G-I, R and T), staphylococcal enterotoxin-like proteins SE*ls* (J-Q, S and U-X) and toxic shock syndrome toxin-1, TSST-1. Out of these, 16 have been structurally determined, the first one, SEB, in 1992 [37], and they all adopt a similar fold, see *figure 4B*.

In this thesis, the superantigen SEH has been studied in complex with MHC class II and its appropriate TCR. An unusual feature is that mice do not show any symptoms when subjected to high levels of SEH, indicating that SEH does not interact with any mouse TCR, which most other superantigens do [38]. SEH is also unusual in that it does not bind any TCRVβ, but upregulates cells expressing TCRVα10 (TRAV27 in the IMGT nomenclature, used from now on) [24, 39].

The advantage for the bacteria in producing superantigens is not fully understood but nature has positively selected for this throughout evolution. SAgs are believed to play a role in early infection rather than late and enhancement of local inflammation may be beneficial for the bacteria in terms of nutrient supply due to the increased blood flow, but basically, 25 years of research has not come up with a satisfactory explanation [8, 40].

The unique features of these molecules have been used by humans to, for instance, engineering SAgs to treat cancers successfully; a Fab, a fragment of a monoclonal antibody together with a chimera of SEA and SEE is in Phase III for clinical trials of different cancer patients showing promising anti-tumour activity [41]. The theory is that the antibody recognizes the tumour cells and that SEA/E recruits and binds to T-cells at the site of the tumour and starts cytotoxic activity to eliminate the tumour cell.

This thesis provide an broadening of the field of superantigen biology in that we prove and explain in detail how a superantigen can interact with the variable alpha domain of the TCR.

Intrinsically Disordered Proteins

The paradigm that the function of a protein lies in one, rigid three-dimensional structure has started to change in the last decade. The first review regarding intrinsically disordered proteins, IDPs, was published in 1988 [42] and since then the field has grown exponentially. These 'unstructured' proteins exist in an ensemble of conformers and are present in all kingdoms of life. 25-30% of eukaryotic proteins are mostly disordered, more than half of the mammalian genome has long stretches (>50amino acids) of disorder and for signalling proteins 70% of the proteins belong to this category (based on DNA *in silico* analysis) [43-45], see *figure 8*. The non-folding is encoded in the amino acid sequence and these are among the most interesting targets for modern protein research.

This class of proteins have forced researchers into new ways of thinking about proteins and the way they interact and function. Previously, protein interactions were explained by the "lock and key" hypothesis formulated by Emil Fischer 1894 [46] where a protein has one set structure with a certain region to which the binding partner fit perfectly. Several models have been proposed to visualize the pluripotent ways of action of IDPs. Fuzziness is a way to emphasize the fluid nature of protein–protein interactions, to explain the "unfoldedness" in complexes [47], to show the scale from static (100% folded) to completely unfolded. The partially unfolded states are the most functionally beneficial due to its adaptability, flexibility and reversibility to the binding of proteins, see *figure 6*.

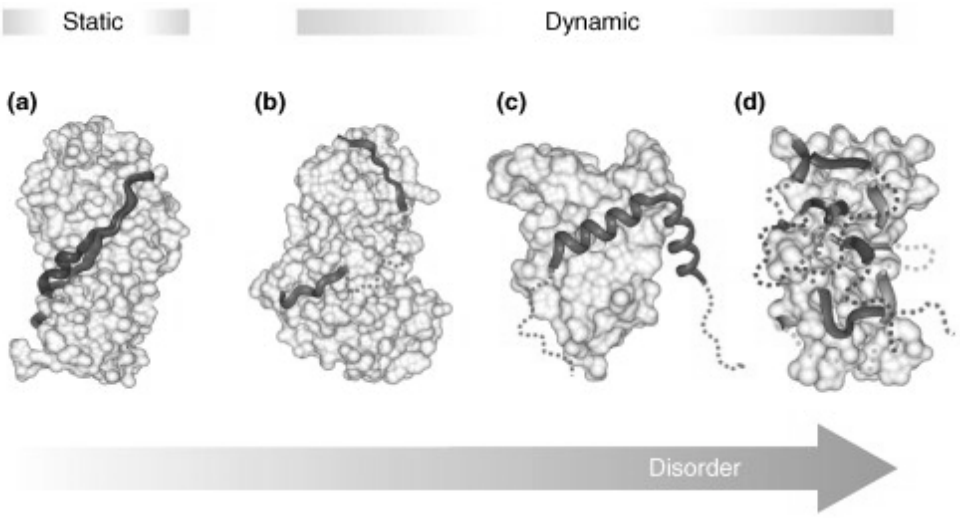


Figure 5. Protein complexes can be all from stable globular proteins to disordered and this range includes structurally well-defined complexes with disordered loops or side chains, complexes with longer segments of disorder, complexes where one partner retains its disorder or where both partners are completely disordered [48, 49]. Note that random coil proteins are outside the range since they cannot make complexes and only exist in denaturing conditions [48]. Printed with permission from [47]

Protein complexes of intrinsically disordered proteins

How does the two proteins decide to make a complex? This raises the question of induced folding or conformational selection; does the protein fold when binding to the proper binding partner or is it the short term structural element that makes the protein recognize its interaction partner? IDPs are, as mentioned, an ensemble of different conformers, in which at some point a peptide may adopt a secondary structure. The ability for a protein to sample a structural element in free form is encoded in the amino acid sequence. These temporary structural elements have many names, e.g. preformed structural element, PSE or molecular recognition features, MORF. They have been shown in several cases to be the functional state of the protein [50-52], and thus, in those cases, we have a conformational selection mechanism [53]. The amount of this structural element in a protein can be measured by following the chemical shift changes for the backbone atoms by NMR spectrum [54, 55]. Induced folding has been confirmed by NMR chemical shift perturbation and relaxation dispersion analysis of the binding of pKID to KIX where binding is followed by folding [56]. Another model for interactions with IDPs is the "polyelectrostatic effect" where the disordered protein has multiple charges and due to the rapid interconversion of several different conformations, an overall "mean electrostatic field" affects the binding affinity rather than a discrete charge in space [48, 57], see *figure 7*.

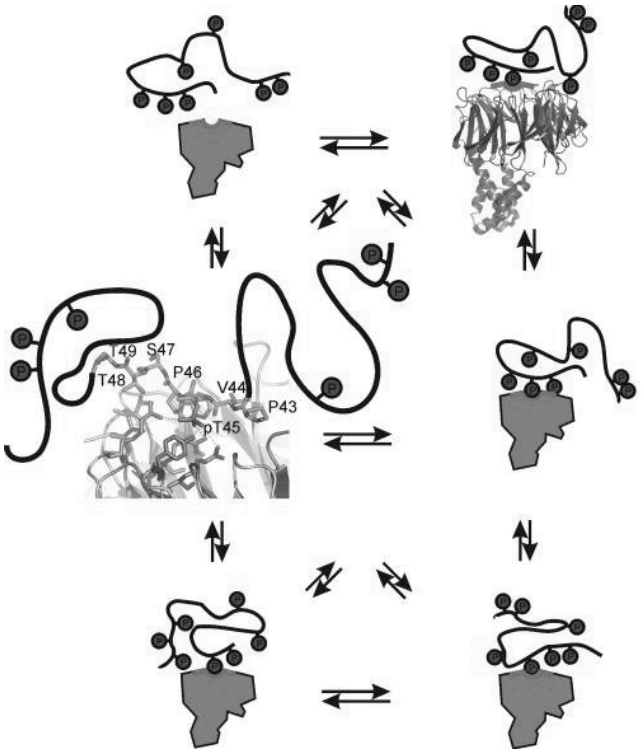


Figure 6. The model polyelectrostatic effect, visualized by Sic1. Upon phosphorylation of its multiple dispersed phosphorylation sites, the disordered cyclin-dependent kinase, CDK, inhibitor Sic1 interacts with a single site on its receptor Cdc4. Sic 1 is shown as a black string and the black dots represents phosphorylations.

NMR analysis show that multiple phosphorylated sites on Sic1 interact with Cdc4 in a dynamic equilibrium with only local ordering around each site. This is an example of protein interaction in a dynamic ensemble of intrinsically disordered states. Printed with permission from [57].

IDPs seem to have several low-minima conformations, contrary to folded globular proteins. In the case of folding-upon binding thermodynamically, there is an entropic cost for ordering which has to be energetically compensated in some way. In some cases, for IDPs, the disorder of another part of the protein is increased upon folding of the interacting surface. Another positive energetic contribution comes from the few, hydrophobic side chains of IDPs, which has unfavourable hydration in the free state and upon binding and/or folding release water – increasing entropy in the system [58].

Bioinformatics studies show a larger-than-average proportion of IDPs in so-called "hub"-proteins, which are central in a protein interaction network and for them the disorder gives the plasticity needed to bind several different partners [48, 58]. This can also be termed moonlighting, and unlike classical cases of globular proteins, these IDPs use the same region or overlapping interaction surfaces for binding [59].

In the cell

IDPs are present in all cellular activities and particularly often found in signalling, translation and transcription activities [60]. Interestingly, IDPs have not an increased preference for interactions with chaperones, which argues that IDPs are different from unfolded or misfolded forms of globular proteins [61]. In cells, the mRNA levels of ordered and disordered protein are the same but transcription of the disordered proteins is slower and their half-lives are shorter due to protease degradation, thus the concentrations of IDPs in the cell tend to be lower [62]. Posttranslational modification is a way to extend the half-life in the cell and IDPs are substrates to twice as many kinases as globular proteins. Phosphorylation is probably the most important and frequently referred-to mechanism of regulation in the cell. It occurs in practically all studied IDPs, which can be reasoned that the three main requirements for posttranslational modifications are fulfilled: appropriate local sequence, exposure of the sequence and ability to adapt to the modifying kinase. In an evolutionary sense, this control of synthesis degradation and modification is related to the major role of IDPs in signalling where it is crucial for a protein to be present in appropriate amount but not for longer than needed [62]. Also, IDPs have longer half-lives in the cell if they have many different interaction partners or are part in larger complexes (e.g. p53) [63].

Disease and IDP

Many diseases have a high abundance of IDPs, see *figure 8*, and the neurodegenerative Alzheimer's (protein Aβ) and Parkinson's disease (protein α-synuclein) are two examples

where the causative agent is an IDP and also the main target for therapy. Often, an interaction partner is the target in therapies for IDP-related diseases, but recently small molecules have targeted the IDP with success for the transcription factor c-Myc, Aβ and an oncogenic fusion protein, EWS-Fli1 [64, 65].

An example where intrinsic disorder is well used is p53. It is a protein of 393 amino acids that has been called the "guardian of the genome" and is inactive in 50% of all cancers. p53 induces or inhibits about 150 effectors involved in regulating the cell cycle, controlling apoptosis and DNA repair, so in other words, p53 is important. It can be divided in four structural and functional regions where some are ordered and some disordered, see *figure 8*. p53 is a homotetramer and on a molecular level it is regulated by a wide array of posttranslational modifications, which are changing with the cellcycle. One single protein making interactions with 150 different molecules, how is that possible? Disorder may be the answer, allowing; overlapping binding surfaces, different conformations at different environmental conditions and overall adaptability, leaving p53 as a true hub-protein [66].

Studying these proteins on an atomic level is not trivial. Up to some years ago, researchers had not been able to characterize them due to lack of suitable biophysical methods. Globular, well-folded proteins can be investigated with (most often) X-ray crystallography and NMR. IDPs, which lack one single structure, require techniques used in solution and for site-specific information, NMR is the strongest technique for examination on an atomic level.

The IDPs investigated in this thesis are cytoplasmic domains from TCR (four) and BCR (two), that all have ITAM motifs and are crucial in the signalling into the cells. All of these domains have previously been shown to be intrinsically disordered by *in silico* methods and circular dichroism, CD [18].

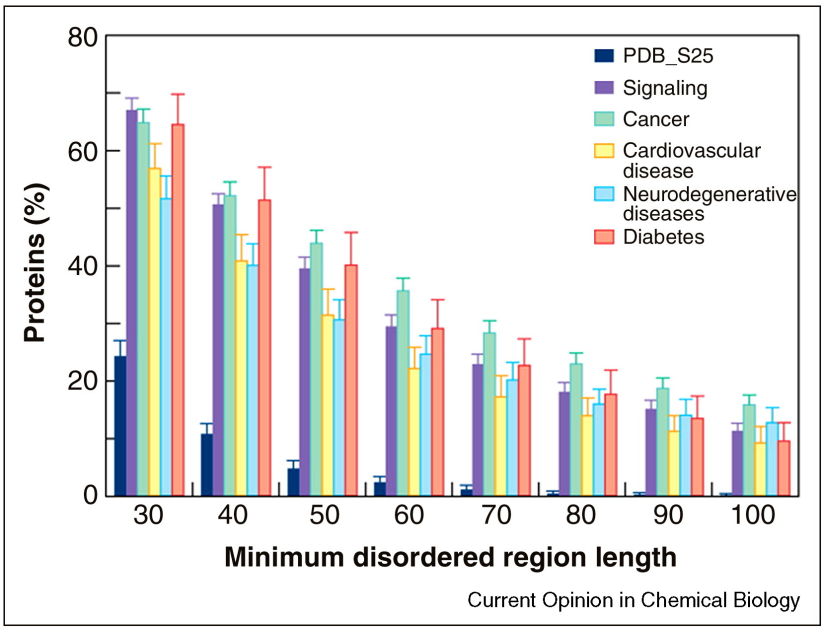


Figure 9. Disorder is overrepresented in proteins involved in signalling and diseases such as cancer, cardiovascular diseases, neurodegenerative diseases and diabetes. A large set of data from each group, ranging from 1786-285, proteins were analysed for disordered lengths and compared to PDB_S25, a subset of 2238 diverse protein chains from PDB [67]. Printed with permission [64].

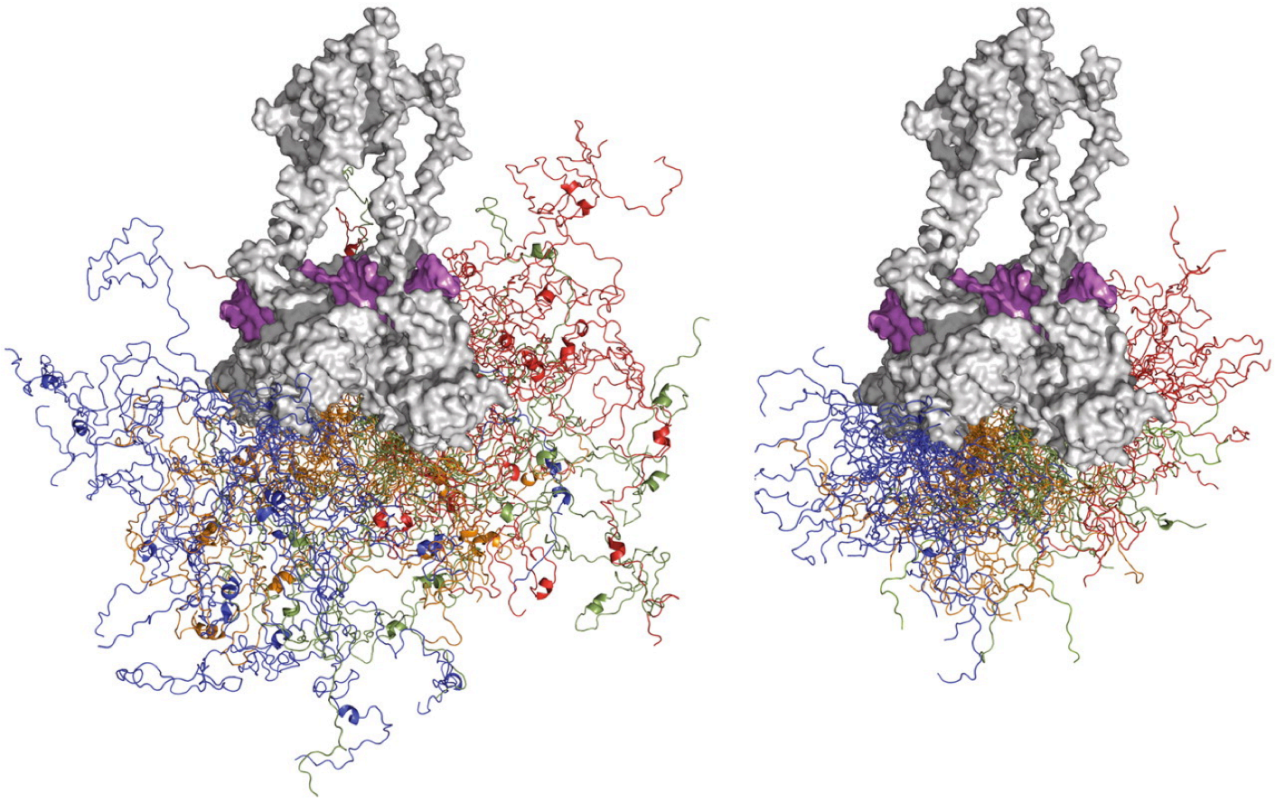


Figure 8. Model of the protein p53. p53 is composed of four subunits, here coloured blue, red, green and yellow. Each which contains well-structured DNA-binding, tetramerization domains as well as a disordered N-terminal transactivation domain. The folded domains were structurally determined using NMR and SAXS. For the intrinsically disordered domain, RDCs and SAXS were used in combination with MD simulations to calculate the average ensemble structure. Printed with permission from [66].

Intracellular domains of TCR

The longest of the TCR signalling domains is CD3ζ, which carries 3 ITAM sequences and has previously been characterized in solution and in detergent micelles. Based on size exclusion chromatography, ζ has been estimated to dimerize at concentrations higher than 10 μM. NMR investigations of ζ in solution have shown no chemical shift differences comparing spectra of concentrations higher and lower than the estimated K_d for the homo dimerization [18, 49], indicating no structure formation or specific interactions between the monomers. In *Paper 1,* we see the same phenomenon where interaction between SIV/Nef and ζ does not give rise to any chemical shift changes in the HSQC spectrum of ζ+Nef compared to free ζ [68].

The other components of TCR are the CD3ε, CD3δ and CD3γ subunits, which pair up in dimers held together on the extracellular side and in the membrane. CD3ε has been characterized the most and similar to ζ, it has been shown to induce an α-helix in the presence of negatively charged POPG/DHPC bicelles.

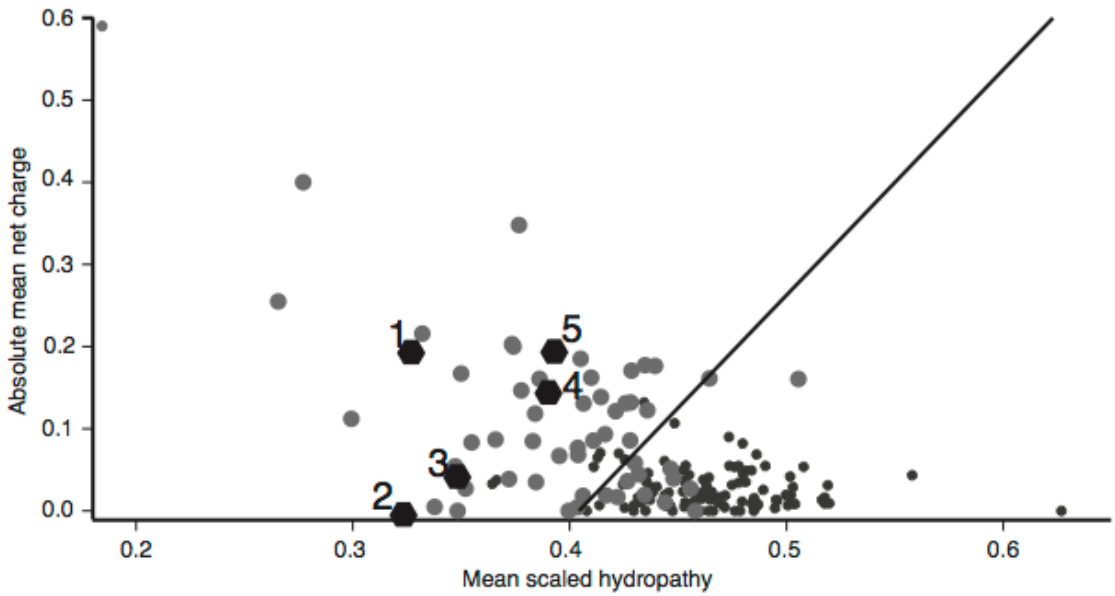


Figure 9. An essential element for IDPs are that they are highly charged and have very few hydrophobic residues, and this is visualized in the Kyte-Doolittle hydropathy plot [69]. Ordered proteins, smaller, black dots, are separated from disordered proteins, larger grey, dots by a line. The plot clearly shows positions of all constructs, black hexagons, 1: CD3ε 2: CD3γ, 3: ζ, 4: CD79a, 5: CD79b. in the disordered region corresponding to a high net charge and low hydrophobicity, characteristic features of IDPs.

Intracellular domains of BCR

In the BCR, CD79a and CD79b pair up as a heterodimers, held together on the extracellular side by a disulphide bridge, similar to the CD3 dimers of TCR, and the signalling subunits are present in the whole maturation process of pro-, pre and mature BCR. CD79a has four tyrosines of which two, Y23 and Y34, are in the ITAM motif. The fourth, Y45, has been found to interact with the BLNK (B-cell linker), important in the signalling cascade [70], indicating non-redundancy in CD79a and CD79b.

In mice, if a cell fails to produce CD79a and CD79b the B cell development does not progress beyond the progenitor stage, leading to immunodeficiency. If only one of them is produced, a partial block at pre-BCR state occurs. Mice with CD79b lacking ITAM as well as tyrosines mutated to phenylalanines in CD79a ITAM show the same phenomenon [71, 72].

Humans lacking, or with defect CD79a and CD79b molecules, show a state called immunodeficiency, i.e. exhibiting less functional B-cells [73]. In lymphomas, several mutations within the ITAM regions of CD79a and CD79b have been found in patient biopsies [74], modifying the function and selection of B-cells and, thus, turning them into tumour cells.

SIGNALLING

Signalling in the TCR and BCR

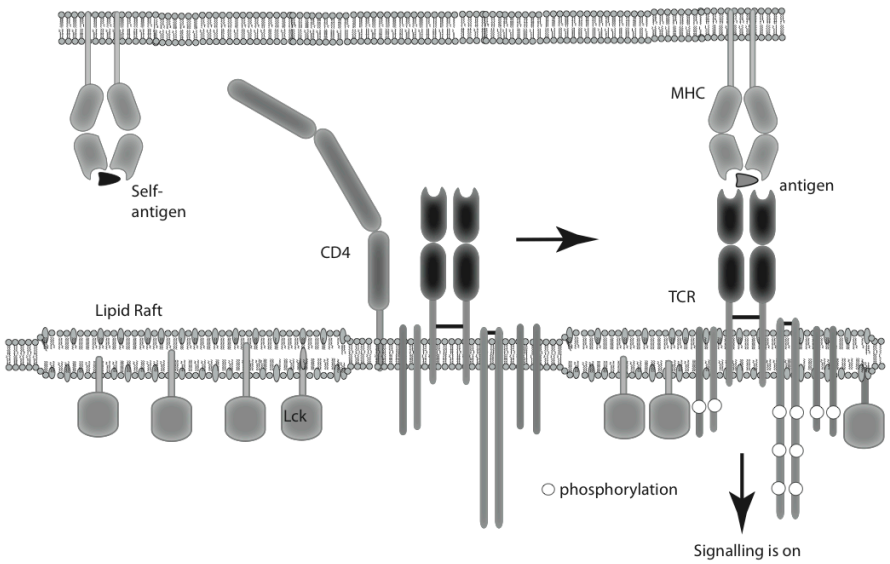
Immune cells are activated by ligand binding of antigen (or superantigen) to the surface bound receptors and this trigger the signalling cascade leading to an immune response. From what is known today, the signalling in the TCR and the BCR are in many ways similar; they both get activated by binding of antigens to the extra cellular domain of the receptor, they utilize co-receptors and the intracellular signalling is initiated by phosphorylation of the two tyrosines in the immune receptor tyrosine-activation motifs, ITAM, (5) on the cytoplasmic domain [75{Kurosaki, 2002 #150]}. In the cell there is a constant phosphorylation by protein tyrosine kinases, PTK, e.g. Lck, Fyn and Lyn and simultaneously a dephosphorylation by phosphatases such as CD45, keeping a low phosphorylation level, called tonic signalling and no downstream activation [76]. Once the receptor is activated, the immunological synapse, IS, is created. IS is the complex antigen-MHC-TCR or antigen-BCR makes up, together with the surrounding membrane and proteins [4, 77]. This includes co-receptors such as CD4, CD45, CD22 and these are either involved in keeping the cells together by recognizing and binding to molecules on the APC or acts by phosphorylating/dephosphorylating molecules in IS [78]. When antigen is bound and IS created, the level of phosphorylation by the first row of kinases on ITAMs increases and these recruit kinases with Src-homology-2 (SH_2) domain. These, in turn, create docking sites for other kinases that have many phosphorylation sites and activate several other proteins such as kinases, adaptor proteins and transcription factors; the signal cascades are on [76, 79].

	TCR	BCR
First kinase	Lck, Fyn	Lyn, Fyn, Blk
Second kinase (SH_2 domain)	ZAP70	Syk
Third signal, kinase (SH_2 domain) or adaptor protein	LAT, SLP-76	BTK, BLNK (SLP-65)
Phosphatase	CD45, CD148	CD19, CD22

Table 1. Examples of the different kinases and phosphatases involved in initial signalling in TCR and BCR. Lyn, Lck, Fyn and Blk are the initial kinases, phosphorylating ITAM tyrosines. This is sensed by ZAP70, T cell receptor Zeta associated protein of 70 kDa and Syk, Spleen tyrosine kinase. These activate SLP-76 Src homology 2 (SH_2) domain–containing leukocyte phosphoprotein of 76 kDa, LAT, Linker for the activation of T cells, SH_2, Src homology 2, BTK, Bruton's tyrosine kinase and BLNK, B-cell linker. The phosphatases are cluster of differentiation, CD molecules. The classification CD has about 350 different molecules in humans and they are providing targets for immune phenotyping of cells. The proteins are cell surface molecules that can be involved in signalling such as receptors or ligands or cell adhesion [80].

The step into the cell and phosphorylation of the ITAM is not clear. Many different techniques have been used and several pieces of evidence make the basis for models of how the triggering of TCR/BCR make a signal through the membrane to phosphorylates the ITAM on the cytoplasmic tails, they can be divided into several groups;

In the membrane/segregation:



The membrane, the outer wall keeping the cell as one, is a fluid ever-changing two-dimensional layer of lipids. These have a long hydrophobic chain facing inward the membrane, and a head group, which may be charged or not. A membrane has 500-1000

Figure 10. A model for membrane involvement in the T-cell activation, including lipid rafts.

different lipids and the composition of them is of great importance for functions such as signalling [81]. The IS is embedded in a so-called lipid raft [4], that is considered to be dynamic, nanometer-sized membrane domains with a particular combination of lipids and proteins, a microenvironment [82]. Lipid rafts have an increase in saturated, charged lipids and cholesterol [83]. One role of cholesterol is to fill in and make the raft more rigid and in general, ordered membranes formed by saturated lipids are thicker and has a different curvature than fluid membranes [84]. The physical change of the membranes around the IS, could serve as a mechanic force, to stretch the transmembrane domains of the TCR/BCR which is transferred in to the cytoplasmic domains causing their conformation change to accommodate the membrane curvature [85]. Removal of cholesterol in the lipid raft has been shown to induce dissociation of activated BCR oligomers [86] [11]. Lipid rafts also include proteins, for example Lck, Lyn and Fyn whose activity would enhance the phosphorylation. On the other hand, certain proteins are excluded from the membrane in the making of the lipid raft, for example, the co-receptor phosphatase CD45 [78, 84]. Upon binding to peptide-MHC, the distance between the cells becomes smaller (about 15nm) than the size of the large ectodomain of CD45 and CD148, these are thus excluded through the fluidity of the membrane [5, 60, 78]. It has been shown that membrane composition changes due to receptor rearrangement with help from enzymes such as flippases, scramblases and phosphoinositide 3-kinase (PI3K) [87].

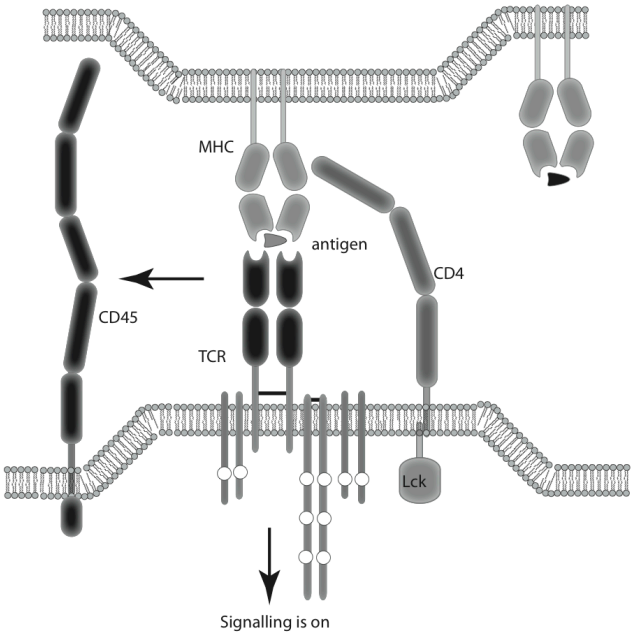


Figure 11. Segregation model where the phosphatases are physically excluded.

Aggregation:

A single antigen cannot start T-cell activation, binding of TCR to an p-MHC results in a clustering of the TCRs and this means action, also called the crosslinking model (CLM). The antigen has to be presented by dimerized or oligomerized p-MHC [88] or if on a APC, together with the co-receptor CD4 [89]. The CD4 co-receptor interacts with MHC on the extracellular side with an angle, almost 90°, and thus, the model of pseudo-dimer proposes a binding of the next TCR to a selfpeptide-MHC due to the CD4 interaction, and making the dual-TCR needed for receptor engagement [21].

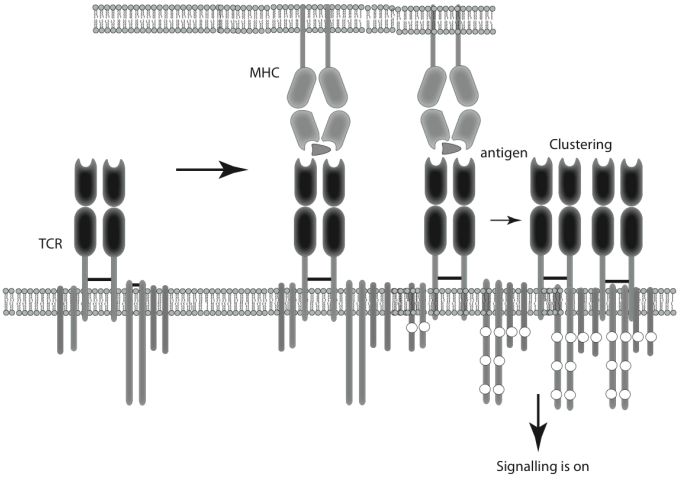


Figure 12. Aggregation of T-cell receptors, initiate the signalling cascade downstream into the T-cell.

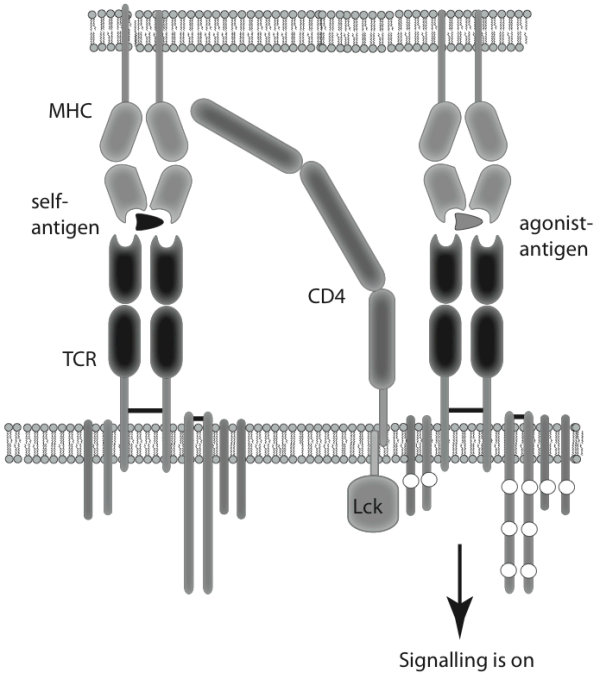


Figure 13. The pseudo-dimer model, making use of self-antigens for starting a T-cell response

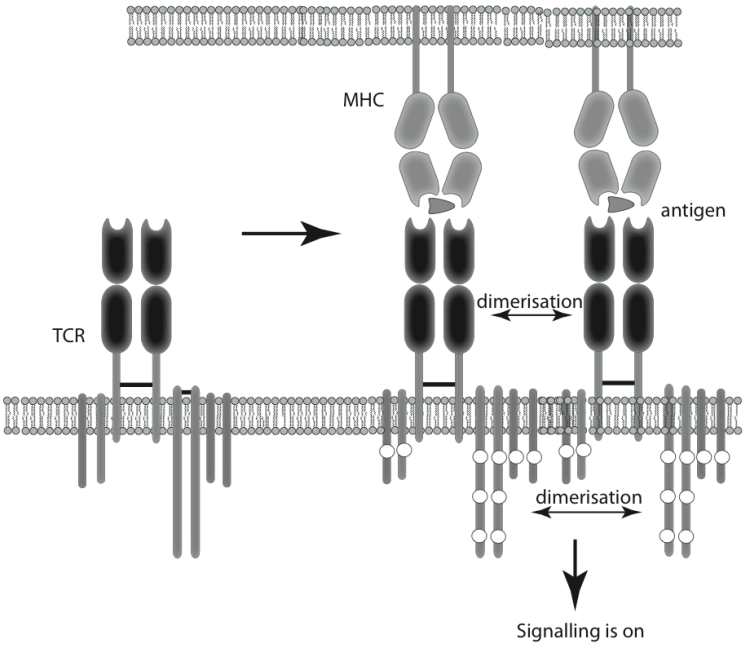


Figure 14. Oligomerization model where the homooligomerisation of the cytoplasmic domains is sufficient for signalling.

A general model is Signalling Chain HOmoOLigomerization, (SCHOOL) according to which, homooligomerization and formation of competent signalling oligomers on the intracellular side, provides the necessary and sufficient event to trigger receptors and induce cell activation. The model can be applied to all receptors using intracellular signalling domains and is based on binding of multivalent antigens that dimerize the receptors [90].

The stoichiometric nature of the resting receptors before activation is not fully clear but most recently, TCR is thought to reside in protein islands of 7-30 TCR [91], which after receptor activation migrate and make up microclusters with islands of, amongst others, the LAT. Both BCR and TCR are associated with the cytoskeleton, particularly, actin filament is important in this island-formation. These microclusters later make up the IS together with antigen or p-MHC [92-94].

In BCR, this is more controversial where monomers as well as oligomers have been stated as the resting-state. [86, 94] The cross-linking model CLM, proposes BCR to be monomers, on the surface of resting B-cells, seen by FRET [86]. In this study they also confirmed oligomerization upon antigen stimulation and they also confirm a 1:1 ratio of CD79a and CD79b in the BCR [86]. The authors propose that upon binding to monomeric surface immobilized antigen, a conformational change in the ectodomain of the BCR, particular the membrane proximal constant domain, makes the BCR oligomerization-competent. When meeting the next antigen-bound BCRs they oligomerize; make up micro clusters that start signalling [22, 95] as thought for TCR. Contrary to this, dissociation activation model, DAM, proposes that inactive BCR forms auto-inhibited BCR oligomers that upon antigen binding turn into clusters of active monomers based on data from quantitative bifluorescence complementation assay (BiFC) [23, 74]. This model also argues that the size difference in antigen that activates BCR and TCR makes CLM not valid for BCR due to its need for precise spacing between the two receptors preferred for signalling.

Conformational change:

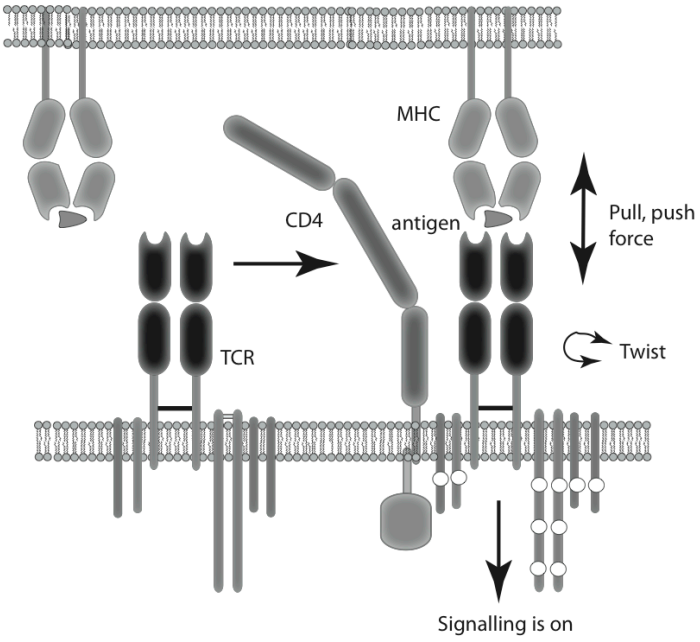


Figure 15. Conformational changes, due to forces between the cells when receptor and MHC bind.

The cause of a conformational change could be a mechanic or physical force between the cells upon receptor binding, signal transduction through rigid domains or physical changes due to rearrangements in the membrane [85]. Pulling, twisting of the TCR has been proposed as a consequence of the antigen binding [78]. The co-receptors CD4 and CD8 interact with MHC class II/class I on the APC and in the T-cell, the intracellular tail of co-receptor CD4 as well as CD8 makes, in presence of zinc, a compactly folded heterodimeric domain with the Src kinase Lck [96]. Upon antigen binding, CD4/CD8 binding to APC may trigger a conformational change resulting in that the Lck gets more exposed for the CD3 cytoplasmic tails. Analysis of several x-ray structures of p-MHC-TCR from pdb (protein data bank) and experiments show that MHC binding induces a small but functionally significant conformational change in the TCR Cα domain at the docking site for CD3ε [97, 98]. Computational modelling shows a high degree of dynamic coupling between the TCRαβ V and C chains, which is dampened upon ligation. This argues that [99, 100] the whole TCR may undergo quaternary change upon activation due to signal transduction based on the rigidity of the components and in that change affect the intracellular signalling units [13, 101].

The role of the intracellular disordered domains:

Another version of the conformational change is the safety model [102]. Shown by NMR data, the protein CD3ε adopts an α-helical structure in POPG/DHPC bicelles, mimicking an acidic membrane. This, aided by a N-terminal basic rich stretch (BRS), buries the two ITAM tyrosines in the membrane and in that, hinders phosphorylation in CD3ε [103] [102]. This safetymodel seems plausible for ζ as well; upon interaction with acidic detergent micelles of LMPG, CD data show increased secondary structure and NMR experiments

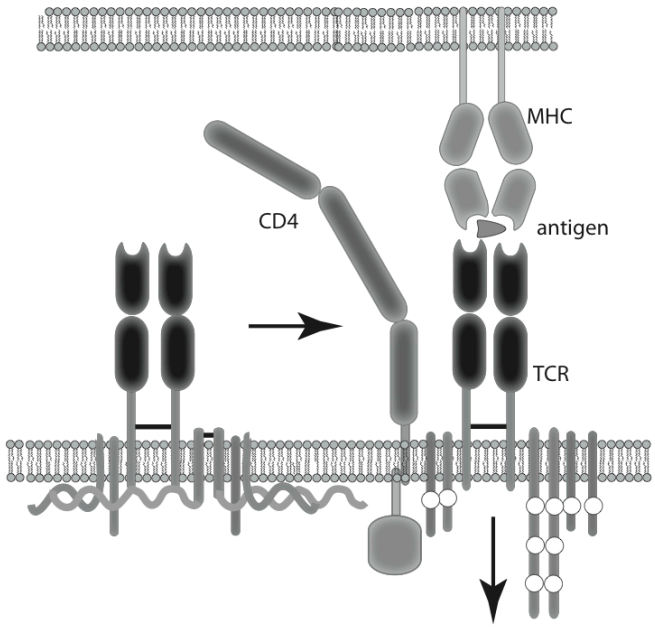


Figure 16. Safety-model where the tyrosines in the ITAM motif get buried in the membrane upon resting cells and antigen binding triggers the release of the domains from the membrane.

confirmation interaction by extensive broadening of a number of resonances, especially in ITAM 2 and 3 [104]. Phosphorylation of ζ was not possible in the presence of detergents [105], in line with the safetymodel. In resting cells, the cytoplasmic tails are bound to the membrane, seen by FRET imaging [106] and engagement of the TCR induces dissociation of ζ from the membrane [107]. Increased local Ca^{2+} concentration, an early sign of activation, also induces dissociation of CD3ε, and ζ due to charges [108].

The signalling cascade moves fast and 4s after antigen presentation the LAT is phosphorylated [109]. The prolongation of the signal is as important as the initiation and whether these two are the same mechanism or related have not been discussed here. Once the cell is active, a continuous presentation of antigens is needed and many different signalling cascades are switched on including internalization of receptors, co-receptors and several feedback mechanisms [110].

Several other receptors in cell with hematopoietic origin use the same set up of intrinsically disordered cytoplasmic domains carrying ITAM sequences for signalling and a extracellular recognition domain, bound together in the membrane by non-covalent transmembrane interactions; 11 major groups with several subgroups, making 28 different receptors, amongst them; Fc Receptors and NK (natural killer)–receptors DAP12 and DAP10 [11, 90]. The motif is found in many other organisms such as *C. elegans* engulfment receptor CED-1 and its *Drosophila* ortholog, Draper, which both have ITAM sequences that get recognized by SH_2 domains [111].

Signalling in cells infected by HIV/SIV

Human immunodeficiency virus, HIV, is the causative agent for acquired immunodeficiency syndrome, AIDS, first discovered by two separate groups in 1983 [7, 112]. HIV and its ortholog simian (higher primate) immunodeficiency virus, SIV, belong to the group of lentivirus in the Retroviridae family. In other words, they are RNA viruses able to replicate in non-dividing cells, hence, slow but very efficient. The genomes are compact, about 9kb and codes for 5 essential proteins (Pol, Gag, Env, Tat and Rev) and 5 non-essentials, amongst them, Nef (Negative factor). HIV infection results in chronic elevation in immune activation and depletion of $CD4^+$ T cells. As the disease progress, the individual develops AIDS and is then more susceptible to all kind of infections [113]. Nef is expressed in abundance in the earl stages of infection and has been demonstrated to enhance viral infectivity of HIV/SIV, downregulation of surface receptors (MHC class I, CD4 and CD28) and modulate immune activation in infected CD4+ T-cells. In studies with monkeys and isolated human cases, individuals deficient in HIV/SIV Nef failed to exhibit disease progression thus, Nef has been suggested to play a vital role in the progression from HIV pathogenesis to AIDS [114].

Full length Nef is 27-35kDa (depending on strain) and all consist of a highly flexible N-terminal anchor domain with a myristylation site and a structured, well-conserved C-terminal domain [115]. The C-terminal domain, called the core domain is the one known for most interactions, about 20 different, and also, this is where ζ is shown to interact [116]. SIV Nef binds ζ at two unique sites SIV Nef interaction domains, SNID, 1 and 2, binding to one of the sites is sufficient to mediate TCR down regulation. It also down regulates CD4 and CD28, this through the AP-2 clathrin-dependent endocytosis. [117-119] SIV Nef has been reported to interact with several proteins in the T cell signalling, including the Src kinases Lck, CD4 and Fyn [120], as well as ζ making ternary complex of ζ-Nef-PTK plausible and a mean for the function of Nef [113]. This binding can be inhibiting or stimulatory for the T-cell. The myristylation modification in the N-terminal of Nef, attracts it to the plasma membrane and IS. Studies have shown that Nef associated with lipid rafts primes the T-cell for activation by increasing the local concentration of TCR. Taken together, Nef is associated with the IS and in the inner leaflet of the plasma membrane it performs its immunomodulatory effects, amongst others, interact with ζ.

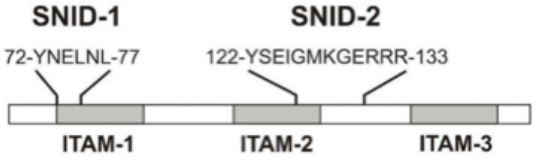


Figure 17. Sequence of the cytoplasmic part of Zeta, showing ITAM motifs and the positions of the two SNID, where Nef interacts. Figure from *paper 1*.

Signalling induced by Superantigens

Superantigens crosslink TCR and MHC independent of the specific sequence of the peptide presented and an extreme amount of T-cells becomes activated. The following signalling cascade is the same as discussed above with kinases and proliferation of the cells. Interestingly in this case is how this extreme activation of T cells is initiated. Superantigens are divided into 5 evolutionary groups based on sequence and they show differences in binding to both MHC class II and TCR.

Superantigen binding to MHC

SAg interactions to MHC have two different sites of binding, one N-terminal with low affinity of K_d 10^{-5} M to the α-chain of MHC and a high affinity, Zn^{2+} dependent site, of K_d 10^{-7} M to the β-chain of MHC [121]. The binding site through MHC β-chain is depending on the backbone of the peptide and uses Zn^{2+} to bridge the C-terminal of SAg to the MHC [31, 34, 122]. SAgs from group I and II bind only the low affinity binding to MHC while group III, and most data points for group IV, interacts with both [121, 123] (SEH has shown MHC interaction using the low affinity site in NMR chemical shift perturbation experiments, data not published). Group V has only been verified binding through the Zn^{2+} to the C-terminal of the SAg [34].

Superantigens binding to TCR

To date, 8 structures of SAg-TCR complexes have been determined using X-ray crystallography and all SAgs bind TCRs using the same cleft of the SAg, except for TSST-1, which has a skewed binding surface see *figure 20*. All complexes but SEH-TCR (*paper III*), show interaction between SAg and the Vβ chain of TCR, the conventional binding site on TCR. No known structure, except SEH-TCR, has included the α-chain of TCR and some Vβ chains include mutations in the interaction surface to enhance the binding in order to achieve crystals [124].

Studies of which T cell alleles get upregulated due to SAg exposure show that some SAgs have a more narrow profile than others. For example, SEA upregulates T cells bearing nine different TCR Vβs, SEB upregulates four, whereas TSST-1 and SEH only upregulates one [8]. Most SAgs upregulate T cells bearing 3-7 different TCR alleles. [8]. The reason for the more general or specific binding lies in the details of the interactions. The interaction can rely on specific side chains of both TCR and the SAg to make e.g. H-bonds, charged surfaces and hydrophobic interactions or it can be more general where the overall shape and

van der Waals forces are the most influential for binding [34, 125, 126]. Detailed examinations of the known structures of complexes SAg-TCR give an insight the different ways SAg bind to TCR.

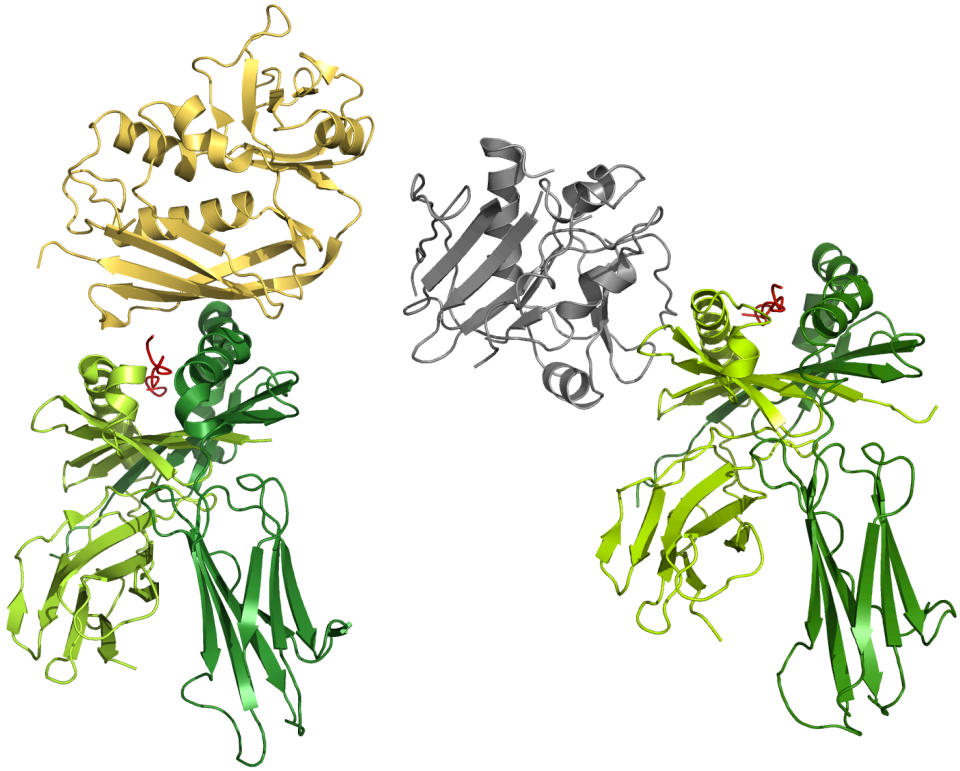


Figure 18. MHC class II binding to SAg in two different ways; high affinity binding to the β-chain of MHC via a Zn^{2+} to SEH, (PDB 1HXY) yellow-SEH, lighter green-MHC α-chain and darker green-MHC β-chain [31] and low-affinity binding to the α-chain of MHC and visualized by SEB-MHC, (PDB 1SEB) grey-SEB, lighter green-MHC α-chain and darker green-MHC β-chain [28].

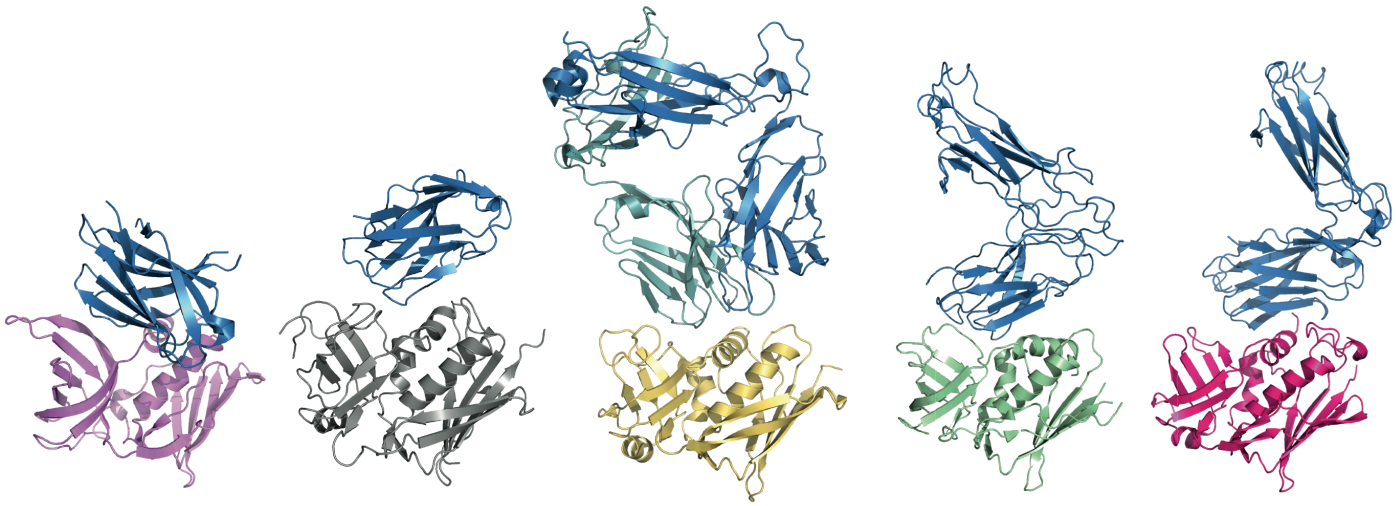


Figure 19. Structure of Superantigens from the 5 evolutionary groups in complex with TCR. TSST-1, purple, bound to hVβ2.1 (3MFG) [127], SEB, grey, bound to mVβ8.2 (1SBB) [128-130], SEH, yellow, to hVα10 (TRAV27/TRVB19) (pdb 2XN9) [1], SpeC, green, bound to hVβ2.1 (1KTK) [131] and SEK, cerise, bound to hVβ5.1 (2NTS) [132]. Vβ is shown in blue and Vα in lighter blue. *Where m denotes mouse and h human construct.*

Detailed analysis of TCR-SAg interactions

TSST-1 recognizes unique amino acids in CDR2 and FR 3, binds only one TCR with no contact between TCR-MHC and is the only member in group I [133]. Group II SAgs are the most general SAgs that interact with several different TCR and through CDR1-2, FR2-3, and HV4 on TCR. The interaction can be described as "conformational-dependent" where the backbone of TCR, rather then the side chains, is of importance for binding and there is a contact between TCR-MHC upon this complex formation [128, 130]. Group III SAg is a diverse group; SEA upregulates cells expressing nine different TCR Vβs but no structural information is available, and SEH only one, through TCR Vα [1]. Group IV SAgs all have a more specific TCR recognition than group II, achieved by interaction with all CDR1-3 and HV4, with contacts both to specific side-chains but also depending on backbone conformation [131]. Group V binds in a more lateral position TCR Vβ. CDR1-2, FR1,3 and, extending into FR4 by an extra long loop which has a unique 15 amino acid insertion, between α_3 - β_8 [132].

Interestingly, the ability to induce emesis (vomit) correlates with their binding mode to TCR. This is due to the existence of an extended loop needed for emesis, present in groups, II and III [134], [34].

The ternary complexes MHC-SAg-TCR

For SEB, the affinity of each and one of the interactions are weak, SEB-TCR, K_d 1,5 x 10^{-4} M, and MHC, K_d 3 x 10^{-4} M, and yet, MHC-SEB-TCR show an affinity of 3 x 10^{-6} M, maybe due to cooperative energetic and stability from TCR Vα in the interaction [135]. SEH interact slightly with Vβ; 6% of the surface in the interface is from Vβ and in the ternary complex a hydrogen bond is created between SEH-Vβ, which is not seen in the binary structure, and thus, not crucial for complex formation [1].

SEH is an atypical superantigen in the group III, seen from phylogenetic tree analysis [34], where it is the furthest away and it differs in that is upregulates T cells dependent on the TCRVα and not TCRVβ, as all other studied superantigens [39]. The upregulation of Vα was further verified by surface plasmon resonance, SPR, analyses where binding was confirmed and a K_d calculated to 4 μM [136], in line with other SAg-TCRVβ complexes, 10^{-4} -10^{-6} M (1-100 μM) [124]. The structure presented in this thesis of the full MHC-SEH-TCR ternary complex and reveals the details of how TCR and MHC are oriented versus each other as well as the specific interaction between superantigen and TCRVα [1]. The overall dual buried surface area is 1369Å, in line with other complexes and the main

interaction points are two hydrogen bonds and a hydrophobic patch. Corresponding residues are found in other SAgs, interacting with TCRVβ and this indicates that SEH is not an outlier, and other SAg may well interact with TCRVα.

Belonging to group III, SEH has potentially two MHC binding sites, something speculated about for SEA [123], the high affinity site is known [31] and a low affinity site shown for a mutant of SEA_{D227A} [123]. This would enable an antigen-MHC/SEH/antigen-MHC/TCR binding, a potential way to increase immune cells response.

It should be noted that there is no overall conformational change in any of the proteins upon ternary complex formation. Another ternary structure, MHC-MAM-TCR was previously determined, and in this complex both Vα and Vβ interact with MAM which is not structurally related to SAg [137].

Affinity maturated SAg

Since the binary complexes between various TCR-SAgs, in low micro molar affinity ($>10^{-6}$ M), crystallization is often been helped by mutations to increase the binding affinity of the complex. This has been done routinely and turned into a specialized tool for investigating affinities and specificities in protein-protein interactions [125]. Interactions depend on more than electrostatic forces, hydrophobicity and van der Waals forces. Mutational studies on SAg-TCR complexes show other factors are involved, such as shape complementarity, water networks, plasticity and cooperativity where 1+1 make more than 2 [124, 125]. Affinity maturated TCR are, besides tools for basic protein-protein research, used for treatment of SAg infection by neutralizing the SAg [8].

Activation by superantigens

Superantigens are very potent and a matter of debate is whether the stronger affinity results in the more immune response [124]. As little as 10^{-12}-10^{-15} g, quantities will induce a measurable activation of the T-cells [34]. (Compare to an ordinary ibuprofen tablet for headache where you take 10^{-1} g, 100mg active substance). SAg mediated TCR signalling follows the conventional activation starting with phosphorylation by Lck on CD3 cytoplasmic tails ITAMs, with exceptions, an alternative way is found [34]. SAgs bind only to MHC class II, which in conventional antigen activated T-cells, binds co-receptor $CD4^+$ but SAg-activated T-cells carries both $CD4^+$ and $CD8^+$, hence, the interaction of co-receptors is of no importance [138]. Studies show that cells lacking, $CD4^+$ produced IL-2 after activation by SAgs, hence, some other pathway is active in SAg engaged T-cells [139]. This was investigated in T-cells lacking Lck where the activation of TCR by SEE (also

group III, as SEH) has shown to proceed through lipid raft-enriched pertussis toxin-insensitive G-protein α-subunit 11, Gα11, and the downstream target phospholipase C-β, PLCβ. This results in translocation of the transcription factors NF-κB to the nucleus and production of IL-2, as in Lck-induced activation [139]. This is another example of the diversity of SAg in activation of T-cells.

In conclusion; nature often use similar designs and means in solving problems of analogous nature, thus, knowledge on the activation of the T-cell and B-cell receptors may give insight in the mechanism of several other MIRR. Understanding on the fundamental ability and mechanism of the disordered proteins are of great need in this emerging field. Superantigens are traditionally thought to activate the TCR through binding to Vβ, but here we show that the Vα may be as justified for interaction, likely using similar binding mechanism. In this thesis we want to shed light on the signalling on the atomic level, using the tool structural biology, which gives very detailed information for each amino acid in the proteins.

STRUCTURAL BIOLOGY

The field of structural biology is about observing and visualizing macromolecules, usually proteins or nucleic acids at an atomic level, meaning each atom and the bonds between them. Traditional light microscopes uses visible light and has a physical limitation in resolution of half of the wavelength used, 400-700nm (4-7x10^{-7} m). That means that structural features smaller than 200nm cannot be observed and the size of single atoms and their bonds are in the order of 0.1-0.5nm (1-5Å), hence, other techniques are needed for these studies.

The two main techniques for structural biology are X-ray crystallography and Nuclear Magnetic Resonance (NMR) and in this thesis both techniques have been used.

NMR is a versatile technique that explores the effect of NMR active nuclei, in biological samples usually ^{1}H, ^{13}C and ^{15}N, in a magnetic field. After excitation by a radiofrequency pulse, the return to thermal equilibrium is recorded and results in the NMR spectra. The molecule is observed in solution and a wide range of dynamics within the molecule may be analysed as well as information about each active nuclei in the protein, for structure calculation and binding studies. NMR was first described in 1938 and was applied on liquids and solids 1946 [140, 141]. 1985 [142] the first protein structure was calculated and since, the technique is still rapidly developing.

X-ray crystallography is a technique that uses an electromagnetic radiation with a wavelength of 10-0.001nm (10^{-8} -10^{-12} m), so-called X-rays, discovered by Carl Gustav Röntgen in 1895. This technique was first used on macromolecules in the early 1900´s. Since then 78800 protein and nucleic acid structures have been structurally determined (as of 3 April 2013), using X-ray crystallography. X-ray crystallography is a very strong technique and the information is very detailed. The prerequisite for X-ray crystallography is a crystal of the protein of interest, which in turn requires a homogenous solution protein of stable fold, implicating that the ever-changing IDPs are not suitable for x-ray experiments. Also, protein complexes with a weak affinity, $K_d > 10^{-6}$ M, usually are very complicated to crystalize but very well suited for NMR experiments where the weaker complexes often give good spectra. For NMR, the challenge is large proteins that give more signal overlap and broader signals. To solve this, isotope labelling (for bio molecular NMR usually ^{2}H, ^{13}C and ^{15}N) of molecules enable multidimensional experiments but yet, there is a size limit.

Protein production

A prerequisite for all structural and biochemical/biophysical work is to have the molecule of interest, the particular protein, pure and if NMR studies, isotopically labelled. Also, the appropriate amount of protein is necessary to get a strong enough signal in your experiment or the proteins to pass the phase transition into crystal state.

There are several ways to produce proteins and one has to take into consideration the organism of origin as well as the production organism for many reasons; posttranslational modifications, the need for chaperones for folding, the balance of different tRNAs and toxicity to the host cell. In this thesis all proteins have been expressed using the "machinery" from the bacterium *Escherichia coli* in different methods;

- Bacterial growth in shaker flask or fermentor producing folded proteins or inclusion bodies – **in vivo** expression.
- Cell free protein synthesis based on *E .coli* extract - **in vitro** expression.

In vivo protein expression

The bacterium *E. coli* is widely used for protein production and since they can harbor and propagate plasmid DNA this is an easy way to introduce new genes and control the protein synthesis of those. The plasmid has a promoter sequence, constant or inducible (such as the T7 and lacpromoters), a ribosome binding site, antibiotic resistance for selection and an open reading frame coding for the protein of interest. For isotope labelling of proteins, the cells are grown in a medium enriched in all or some of the NMR-active nuclei ^{2}H, ^{13}C and ^{15}N, which may result in lowered bacterial growth.

SEH is a toxic protein that has genetically been modified with addition of a signal peptide for transport to the periplasmatic space between the inner and outer membrane of the bacterium, where the amount of other proteins is much lower than in the cytoplasm making the purification easier. The promotor for SEH is constantly on, producing high amounts of SEH [31].

TCR and MHC class II are heterodimeric proteins that natively attach its counterpart in the membrane. The constructs used in this thesis are modified to lack the transmembrane domain. Because of it's hydrophobicity, detergents would need to be included in purification to solubilize this part of the protein [143]. In the TCR monomers, a disulphide bridge has been introduced to keep the heterodimer together (Cys 158α and Cys 171β) [144]. However, the TCR and MHC chains are produced separately and do not fold correctly upon translation, but form insoluble inclusion bodies. The benefit of this is a lot of

polypeptide is being produced and the inclusion bodies are easy to purify [145]. Refolding of these and at the same time, forming disulphide bonds within and between the two monomers (for TCR), making the correct heterodimer is not trivial. The correct redox conditions, slow adjustment from high to low concentration of denaturant by dialysis or dilution, right temperature and additives of small molecules such as the peptide present for MHC refolding are crucial [144-146].

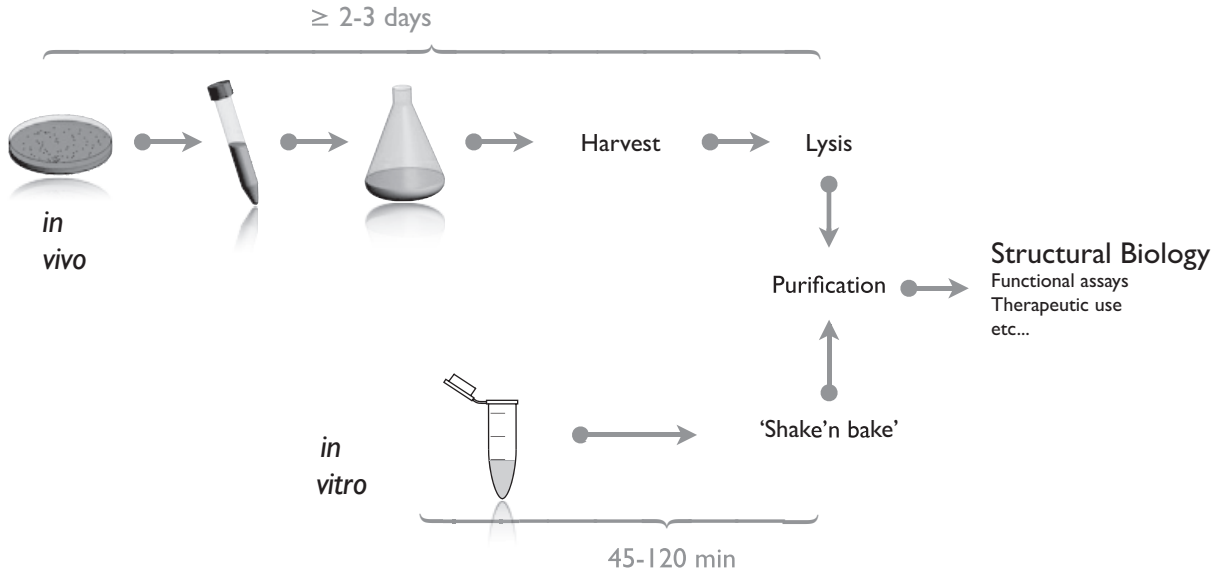


Figure 20. ***In vivo*** **vs.** ***in vitro*** **(CFPS) protein productions**. Preparations: **In vivo**; Cloning of cDNA for the protein of interest into an appropriate vector, transformation into host cell and selection for positive clones. **CFPS**; cultivation and preparation of cell extract, expression and purification of DNA for the protein of interest, making of stock solutions of additives such as nucleotides, T7 polymerase, tRNAs, amino acids, DTT and buffer ions.

The production in *in vivo* includes scale up and growth of a positive clone in shaker flask or fermentor (more control over the process and often better protein yields), lysis of cells and eventually, purification. This takes about 2-3 days where the protein production in CFPS is fast and require accurate pipetting and mixing of all ingredients in one tube, incubation for 45 min-2h followed by purification. Preparations can be performed a few times/year and stocks kept in freezer/fridge.

Cell-free protein synthesis (CFPS)

Nature has under a very long time developed the system of reading information from the DNA code, transcribing it into RNA and then translating that information into proteins which in turn perform the function in all living cells. This is a beautiful and intricate system very well regulated and the basis for all life. Cell-free protein synthesis is a fairly new technique that is being developed at the Swedish NMR Centre amongst other places. In

cell-free protein synthesis, bounding cell membranes have been taken away and consequently, the production of proteins is performed in an open system. The open nature of the system allows the biophysical conditions of expression to be tailored to the synthesized protein.

The cells from *E. coli,* wheat germ, or rabbit reticulocyte are lysed, purified from cellular DNA, membranes and larger components and the remaining crude cell extract contains ribosomes, elongation factors, metabolic enzymes and many more components needed for protein production. This is put into a test-tube and DNA coding for the protein of interest, energy substrate and other components are added, see *figure 21*. Production of proteins can be performed under special conditions such as different pH, redox potential, with RNAse and protease inhibitors as well as co-factors or ligands present. The system is very adaptable and production of several proteins at the same time can be accomplished [147, 148]. In our system, production of membrane proteins with additions of detergents or nano-discs is possible and in a biased screen of available constructs, 37/38 proteins expressed well for structural studies [149]. The set-up can be of either so-called batch mode, where all components are added together in a test-tube or continuous-exchange where the ribosomes and produced proteins are on one side of a membrane and a solution of the amino acids, energy substrates and by-products are continuously exchanged. For a more extensive review, see Carlson el at [148].

At the Swedish NMR Centre, we have developed a batch mode CFPS system based on an *E. coli* extract. [150] The benefits compared to continuous-exchange are simplicity of setup and scale-up. Proteins are expressed for 45 min - 2h and thereafter purified, if so preferred. In CFPS, the only protein produced is the one corresponding to the DNA added to the CPFS mix, hence, it should be possible to run NMR experiments directly on the cell-free reaction (ionic strength and viscosity need to be considered). IDPs have properties proven favourable for production in CFPS, such as high net charge and low hydrophobicity [151], and taken together with the short production time in presence of protease inhibitors, IDPs are well suited for CFPS.

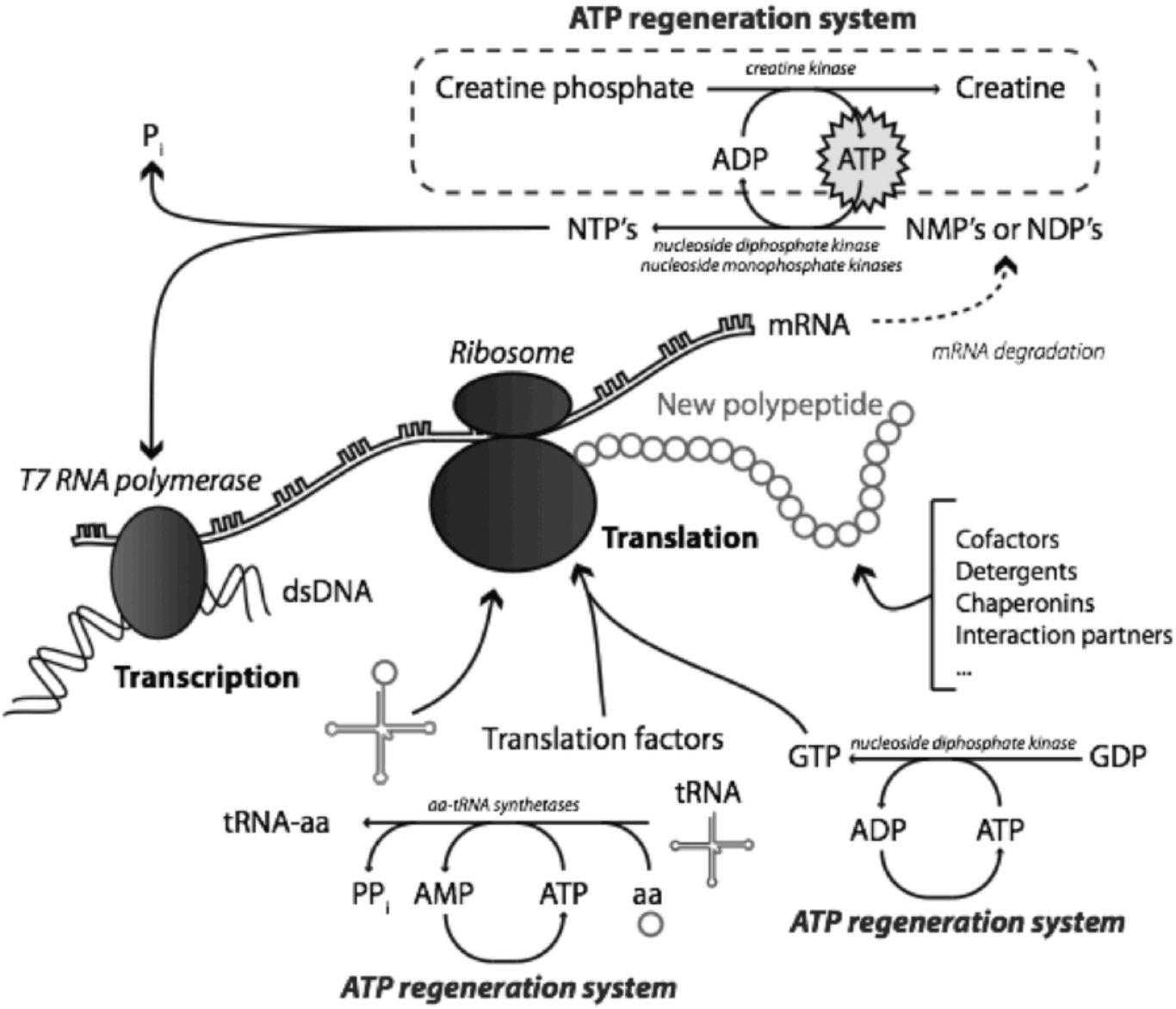


Figure 21, Scheme of the cell-free protein synthesis. In this system, the added genes (dsDNA) are behind a T7 promotor controlled by the T7 polymerase, for specificity and efficiency. The ribosome is from the extract and it uses tRNA, amino acids and ATP to translate the newly synthesised mRNA into the new polypeptide. As energy source, creatine kinase and creatine phosphate are added to regenerate ATP and since creatine phosphate is not endogenous to *E. coli*, the phosphatase activity is low. The concentration of Mg^{2+} is of great importance for amongst other activities, the function of the ribosome and varies from extract to extract and thus, has to be adjusted for each extract batch. RNAse inhibitors, nucleotides, tRNAs and, in total, about 25 different molecules are added to the extract. Increased levels of serine and glutamine are needed due to degradation by serine deaminases and glutaminases present in the *E. coli* extract [150](+A. Pedersen, unpublished).

For NMR, the labelling of amino acids does not affect the yields and time of protein production, which often is the case for *in vivo* production where also scrambling of the labelled atoms is a problem. The amino acids are added as ready-made building blocks, that makes selective labelling and addition of specifically modified amino acids possible [152]. This may be of great importance for large proteins or IDPs. Limiting for CFPS is production of proteins requiring post-translational modifications, disulphide bonds and proteins needing folding assistance. The technique is constantly being improved and at SNC approaches such as metabolomics, experimental design, modifications of the *E. coli* strain as well as empirically trying out different additives such as chaperones are used.

Biophysical techniques

X-Ray Crystallography

Proteins of a fairly rigid structure will in favourable conditions form crystals and this is the prerequisite for X-ray crystallography. The most common technique to use is called vapour diffusion or hanging/sitting drop where the protein solution slowly evaporates and the protein phase transitions into a crystalline form, first a nucleus followed by growth of a crystal. This is often the bottleneck of structure determination and many factors may influence the crystal growth; pH, temperature, ionic strength, protein concentration and homogeneity. The smallest unit in a crystal is the asymmetric unit which is translated in space to make up the unit cells, which build up the crystal. The crystal is exposed to short waves of light and by the electrons in the crystal; the light will scatter in several different directions. Most photons will be lost due to destructive interference, and those in constructive interference will amplify and the resulting collected reflections holds information about the special relationship between electrons in the crystal. One single molecule would give too weak signal so the crystal is needed to magnify the signal from each and one of the electrons in the protein. The crystal is rotated to give data from 360° or less depending on the symmetry within the unit cell in the crystal. To calculate the electron density map knowledge on the phase of the diffracted signals is needed, but the collected data do not contain information about the phases, which is known as the *phase problem*. If a similar structure exists, homologous replacement can be used. Otherwise different methods where heavy atoms are included in the crystal are needed, to get the correct phase. The data is processed and in the resulting electron density map over the electron clouds in the

asymmetric unit, a model of the protein can be built [153, 154]. For a comprehensive text on X-ray crystallography, see Rhodes [154].

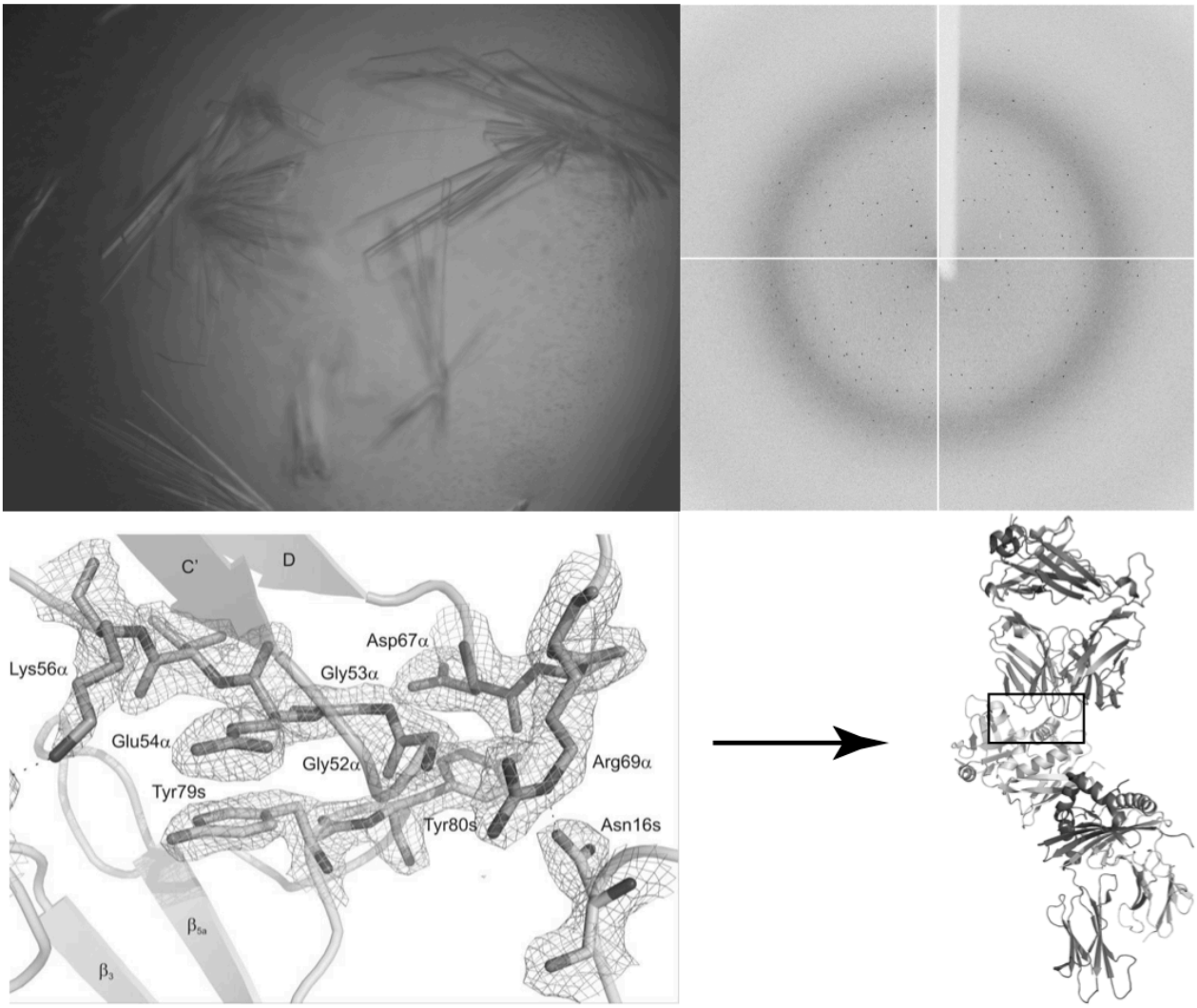


Figure 22. The crystallography pathway. A crystal, here visualized by a drop full of SEH-TCR-MHC class II with bound peptide crystals, is carefully placed in the X-ray beam and the diffraction pattern after X-ray exposure is recorded. The black little dots on the detector are the reflections that are indexed and integrated over all degrees of data collection and then processed. The resulting electron density is shown as a mesh around the model of the peptide chain threaded through the data. The final calculated model of the protein in the crystal is shown last, with a box indicating the exact position of density map shown. Picture adapted from *paper III* [1].

NMR Spectroscopy

The building blocks of molecules, atoms, consist of a nucleus with surrounding electrons. A nucleus has some fundamental properties such as mass, charge and spin where the first two may be understood from experience but the third, the spin, is a magnetic property of the particle, forming the basis for NMR spectroscopy. Each nucleus has a particular gyromagnetic ratio, γ, which in a magnetic field, B_0, has the spinning frequency $\omega = -\gamma B_0$.

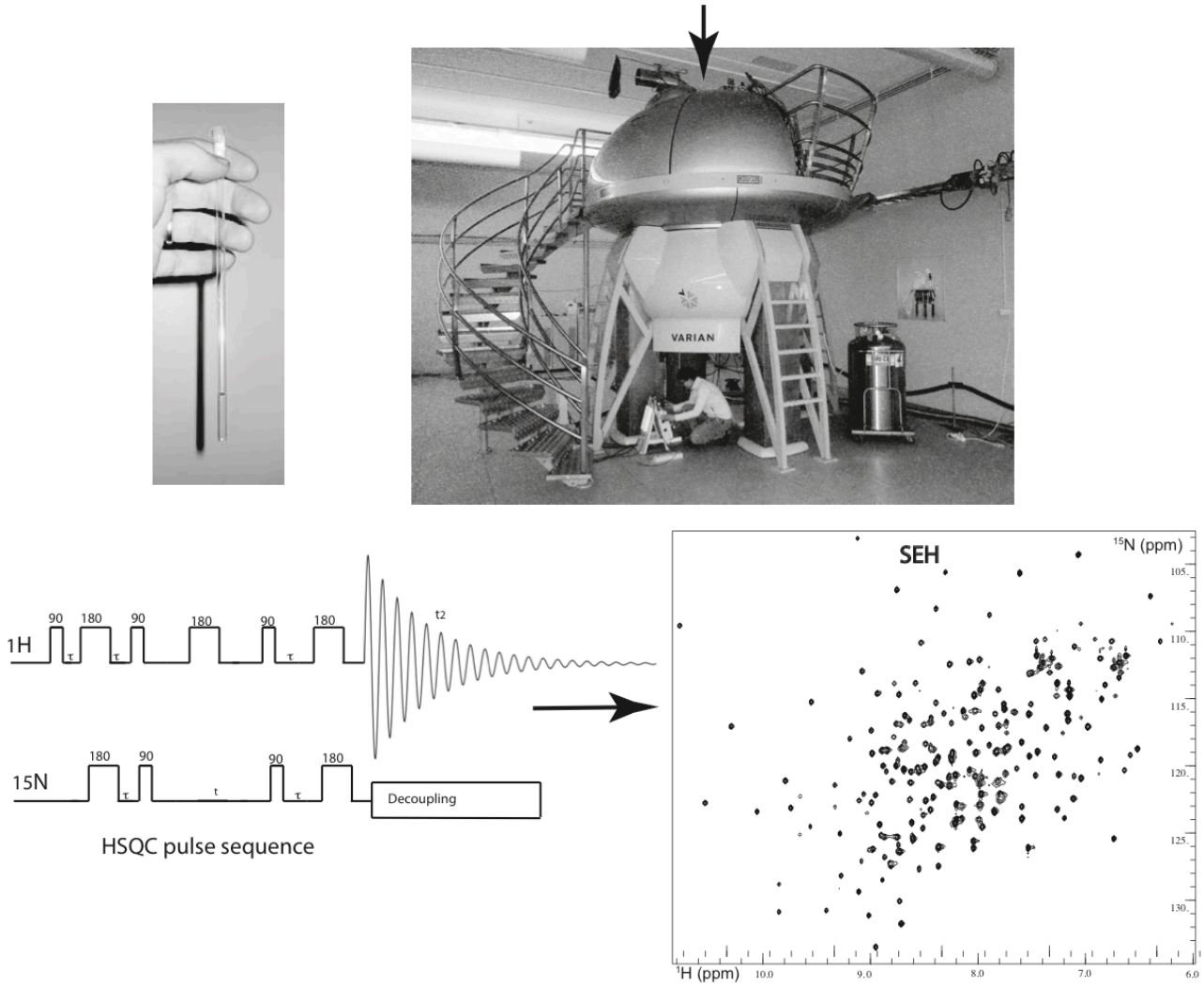


Figure 23. The NMR pathway. A solution with the purified, labelled protein in a thin, sealed, NMR tube is placed into the magnet, here visualized by the 900 MHz NMR spectrometer, equipped with a coldprobe, at Swedish NMR centre. A set of rf pulses are given to the sample in the magnet and the signal of the nuclei returning to equilibrium is recorded and processed, resulting in a spectrum, here visualized by a 2D HSQC spectrum of SEH.

In NMR spectroscopy, the interaction between nuclear spins with spin =1/2 (other spin exist but is not discussed here) and a strong magnetic field is studied. Each nucleus feels a slight difference in the local magnetic field due to an induced, small magnetic field in the electron clouds created by the external magnetic field. The nuclei feel the surrounding electron cloud and this makes a slight difference to the spinning frequency of each nuclei and is the basis for the chemical shift [155].

NMR experiments are a set of radio frequency, rf, pulses, so-called *pulse sequences* applied to the molecule solution. The nuclei, in resonance, will absorb the energy from the rf pulse, transfer the energy to the neighbouring nuclei through bonds or space. In the end the energy is emitted during the return to thermal equilibrium and recorded and transformed from time to frequency domain. The resulting spectrum includes information of the frequencies of each and one of the nuclei involved in that experiment. 2D and 3D experiments help to resolve one peak from another in the spectra and the particular shift for each nucleus is achieved and necessary for future assignment (identification). NMR measurements are in practice the difference in population of two energy states, which is very low and hence, strong magnets and high concentrations of samples and many iterations of the experiments are needed to enhance the signal [156]. For a more extensive text on NMR, see Levitt [156].

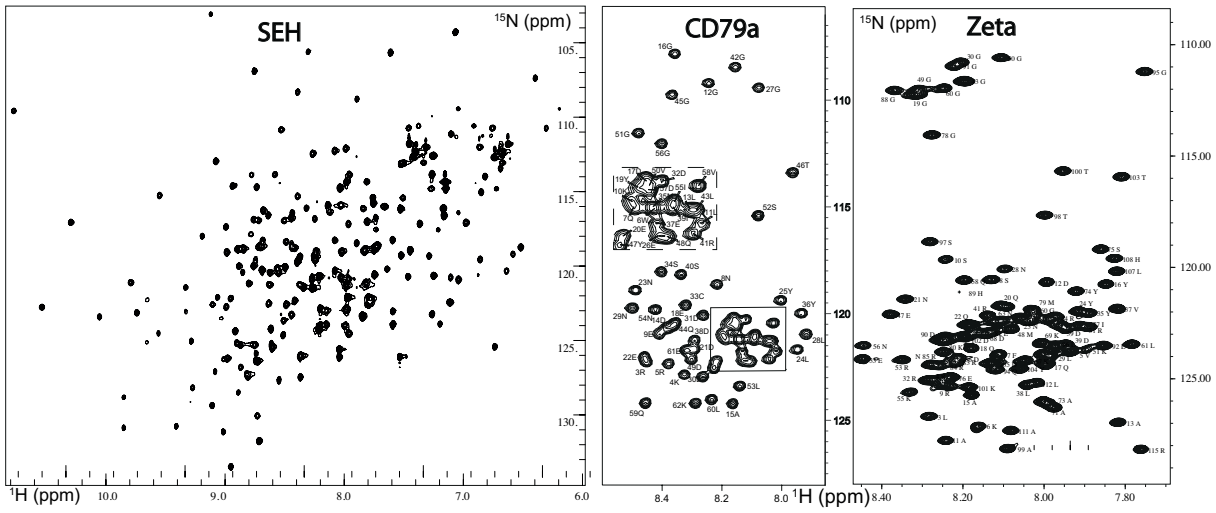


Figure 24. N^{15}HSQC of globular 25kDa SEH, of IDP Zeta of 13kDa and projected HNCO of IDP 7kDa CD79a. In a globular protein such as SEH, the dispersion of peaks is wide but due to slower tumbling rate for larger molecules it results in broader peaks as well as and the increased number of residues, there is a problem with peak overlap. For IDPs, most residues feel about the same magnetic environment and the chemical shift difference nuclei is small, which may result in overlapping peaks. Solution to this is multidimensional NMR, here represented in a 2D N^{15}HSQC. Figures from [104, 157]} and *paper IV*.

Samples are largely unaffected by the irradiation of the rf pulses and the strong magnetic field and can be used over again and even proteins inside a living cell has been analysed. The concentration of proteins in the cell is about 300mg/ml protein, a very viscous and compact solution, not at all the same as in our experiments; mostly water, some buffer- and salt-molecules with the few mg/ml protein. Interestingly, *In Cell* NMR, has shown that some IDPs have the same unfolded state in a cell as in the test-tube [158].

NMR is the main technique to get detailed information about IDPs but also to verify that what is seen in a crystal structure actually is the case in solution, as well.

Assignment

Each cross peak has information about the chemical shift in two or more dimensions where each dimension is one type of nucleus. Combining chemical shift information from several multidimensional NMR experiments to spin systems i.e. residues, and identifying which residues these originates from, is called assignment. An analogy is solving a large puzzle of 3 dimensions or, often more. A set of different experiments are recorded and each cross peak appears in several of the spectra, making it possible to match and assign each and one of them. NMR experiments can show through-bonds (e.g HNCO, HNcaCO, HSQC) which are used for assignment or through-space interactions, Nuclear Overhauser Effect (NOESY) that show special relation between nuclei.

The possibility to solve this puzzle of assignment automatically has developed, with computer power and interest from the community. Sometimes though, the human eye needs to watch and solve ambiguous peaks and often in proteins, there are several conformations of one residue or certain regions that may show very different line widths making the peaks not detected by the software.

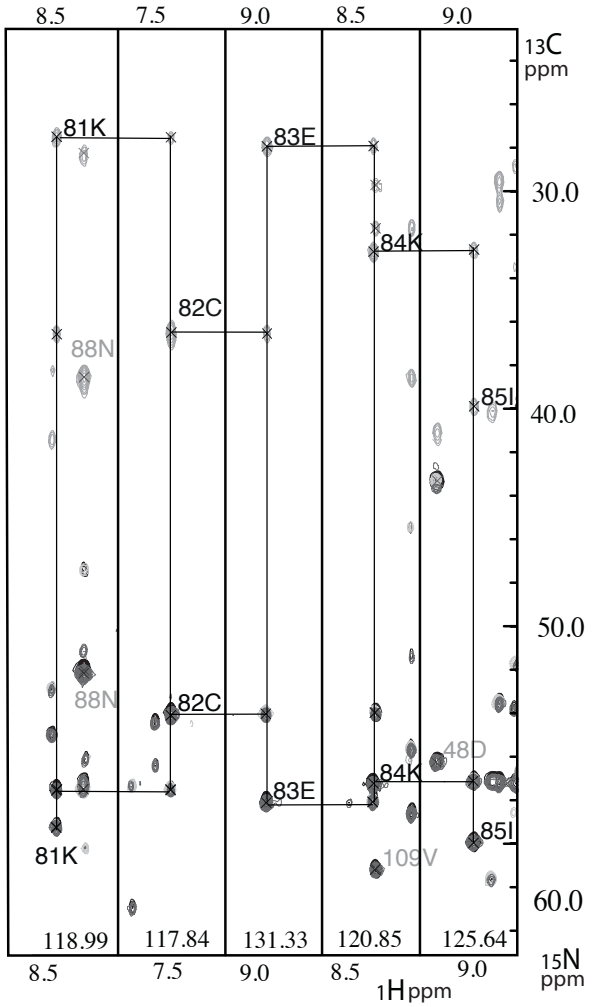


Figure 25. Connectivity in strips of HNCaCb (grey and black) and HNCa (dark grey) spectra from SEH assignment of residue 81-85. On the bottom the ^{1}H dimension, the vertical scale has ^{13}C and inside the boxes are the third dimension, ^{15}N, shown. Cα shift are the lower, around 50-60 ppm and the Cβ the upper, at 25-40 ppm. Both experiments give peaks for i and i-1, which makes the connections possible. Most often, the stronger peak belongs to i and the weaker to i-1, and several experiments show the same crosspeak. For example, HNCaCb and HNCa both show Cα i and i-1, whereas HNCoCa is an experiment showing only i-1).

Automation in assignment and fast spectra recording

As mentioned above, several programs have been developed for automatic peak detection and assignment [159]. There are also several techniques to speed up the actual acquisition of NMR experiments [160] [161] [162]. Traditionally, recording of a full set for backbone assignment takes several days, depending on protein concentration and stability and the sensitivity of the NMR spectrometer. Presented in *Paper IV* is a method for fast recording of NMR experiments and simultaneous assignment of the spectra; Targeted Acquisition NMR Spectroscopy, TANSY. The acquisition of spectra is fast, both in that non-uniformed sampling (NUS) [163] is used for recoding of only a few percentage of a full dataset, the experiments are using the fast-pulsing BEST-TROSY experiments [162] and it is done in a iterative manner with continuous processing and analysis, called Targeted Acquisition, TA [164]. First a HNCO experiment is recorded to have a starting spectrum where each N-H bond from the backbone is represented as one crosspeak. Thereafter, for

backbone assignment, a set of 3D experiments is recorded simultaneously with some increments of each experiment at a time and constantly, the program processes the data and tries to assign the whole sequence using AutoAssign [159]. Continuously, statistics of the progress is given and this enables adjustments in amount of iterations on certain experiments to make a balanced progress with sufficient data in all experiments, see *figure 26*.

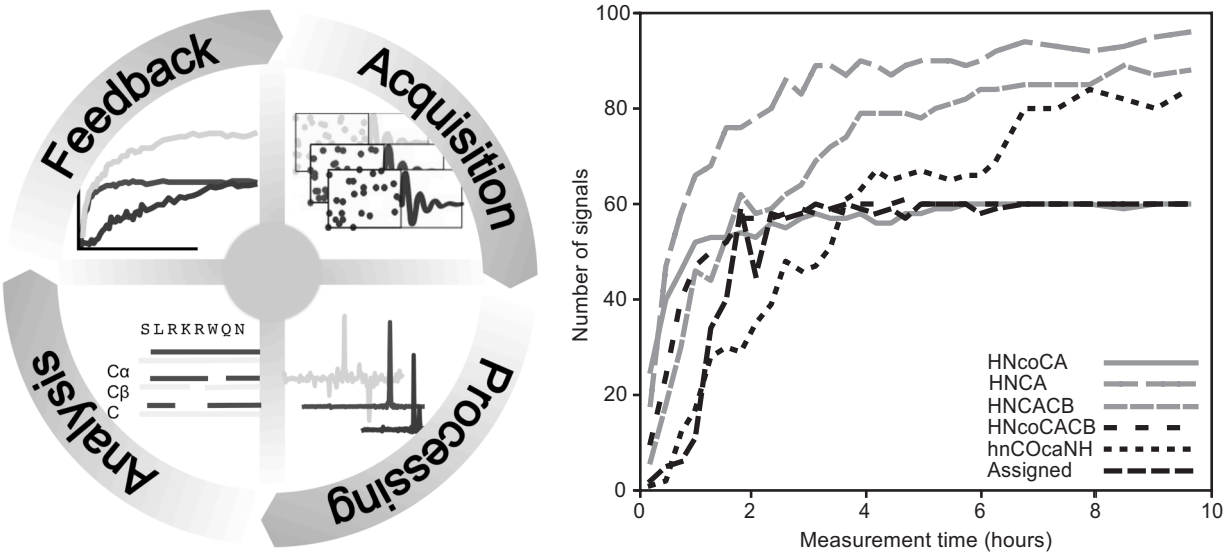


Figure 26. The flowchart of TANSY and progress of TA. TANSY uses iterative acquisition of NMR data, processing and analysis. Each cycle results in validation of the assignment and the process is on going as long as the assignment is improving. The graph shows number of detected peaks in individual experiments versus time in a TANSY set up of a CD79a sample of 120 μM. Figures from *paper IV, with permission*.

Secondary Chemical Shift (SCS)

Chemical shifts can be used to probe full or partial secondary structure. Here, the measured chemical shift for the backbone $^{13}C\alpha$, $^{13}C\beta$ and ^{13}CO are compared with either a amino-acid-specific "random coil" shift from a database based [165, 166] or chemical shifts from denatured protein, i.e. use of 8M UREA or 6M Gua-HCl, guanidinium hydrochloride [167]. This is often used to confirm a calculated structure or for IDPs to visualize structural tendencies, i.e the presence of a secondary structure in a portion of the proteins, e.g. MORF. To enhance the weak presence of an α-helix, Trifluoroethanol, TFE, known to favour states where the backbone amide groups form intra-molecular hydrogen bonds, such as in the α-helix is added. This if up to 20% is used. By this addition, a weak presence of a α-helix will be enhanced and detectable [168, 169].

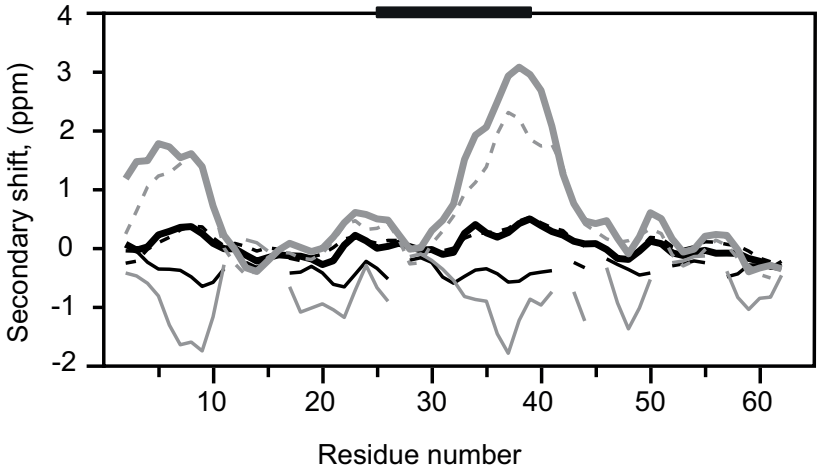


Figure 27. Secondary chemical shifts (SCS) for CD79a, ^{13}Cα (bold lines) ^{13}Cβ (thin lines) and ^{13}CO (dashed lines). Resonances were calculated by subtracting intrinsic random coil chemical shifts [170, 171], measured in 6M urea, from the corresponding values obtained for "native" (black lines) and "20% TFE" (grey lines) states. For presentation the values were smoothed by three-residue running average. The position of the ITAM motif is visualized by a thick black line above the graph. Adapted from *paper IV*, with permission.

Interactions investigated by NMR

Chemical Shift Perturbation

If assignment is known, changes in shift for ^{15}N-^{1}H cross peaks for the protein in solution with the binding partner, are good markers for interaction or subsequent effects in the protein. 2D ^{15}NHSQC or similar spectrum is recorded and the binding partner added followed by recording of a new spectra. The area in the proteins affected by interaction, either by direct binding, being close in space or maybe by secondary interaction such as conformational change, will show a chemical shift difference for the significant cross peaks.

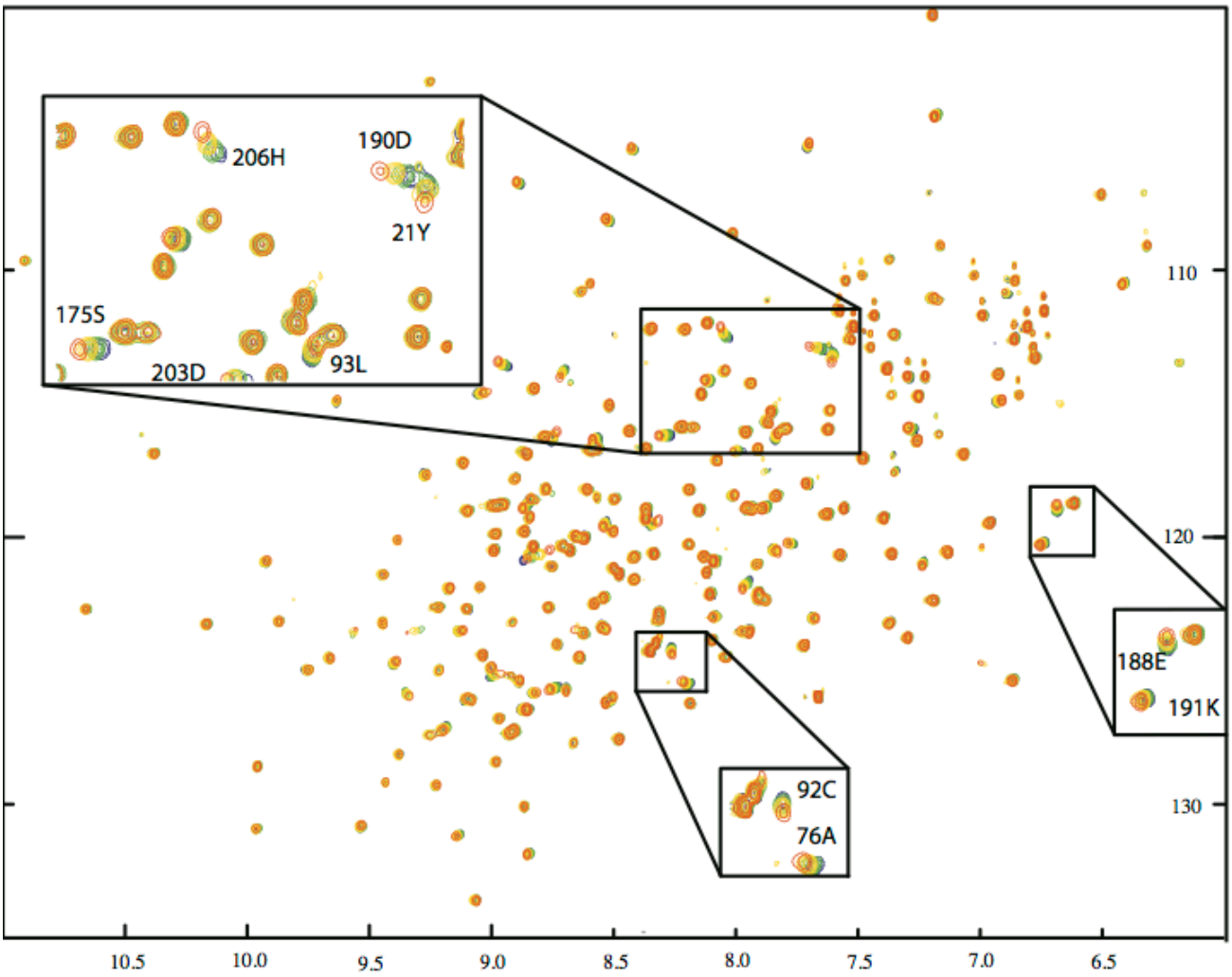


Figure 28. Chemical shift perturbation of SEH upon titration with increasing amount of TCR. (Blue – only SEH present and as the amount of TCR increases, 1:20 (TCR:SEH) green, 1:10 yellow and 1:6 red, some peaks are shifting and/or showing change in intensity. For the interaction SEH-TCR we wanted to confirm the binding site seen in the X-ray structure as well as get additional information about the nature of the interaction. Figure adapted from [1] (*paper 3*), with permission.

Exchange

Proteins are dynamic and constantly atoms are moving in the proteins as a part of life and protein function. This can conformational changes, making and breaking of complexes or backbone chemical exchange processes. Dynamic processes cover a broad range of timescales and different NMR experiment can address distinctive timescales [155, 156, 172, 173]. Folding and interactions are often in the µs – ms and can be mapped using chemical shift perturbations.

If the complex of two molecules has a long lifetime (relative to $1/\Delta\omega$) two distinct peaks will appear, one for the free and one for the bound form of the labelled protein. If, on the other hand, the complex is short lived, the system is in fast exchange and the average peak position is what will be seen in the spectra. Often a short-lived complex is weaker than a long-lived and by analysing the relative change in chemical shift and peak intensities, extraction of dissociation constant, K_d, can be made and often rate constants K_{on} and/or K_{off} can be obtained.

$K_{ex} >> (\omega_A - \omega_{AB})$ = fast exchange

$K_{ex} << (\omega_A - \omega_{AB})$ = slow exchange,

$K_{ex} \approx (\omega_A - \omega_{AB})$ = intermediate exchange.

Other biophysical techniques for investigations of intrinsically disordered proteins

Computational methods

In the field of IDPs, where the unfoldedness is encoded in the sequence of amino acids and with experimental techniques to identify IDPs have allowed for development of many *in silico* tool to find, organize and explore these proteins. Analysis of primary sequence is used in the Charge-Hydropathy, CH, plot (*figure 9*) where few hydrophobic vs. hydrophilic residues i.e. is a first sign of disorder is visualized [69]. Other predictors calculate energies and probabilities of folding based on pairwise inter-residue interactions, e.g. IUPred [174]. More sophisticated predictors, e.g. PONDR-VLXT use a neural network that predicts disorder based on training sets and gives a score for each amino acid [174, 175]. Structural elements such as MORF and PSEs, as well as docking sites for interactions, linear motifs, LM, can be predicted and some servers combine several predictors into so-called Meta servers, such as MetaPrDOS [50, 176].

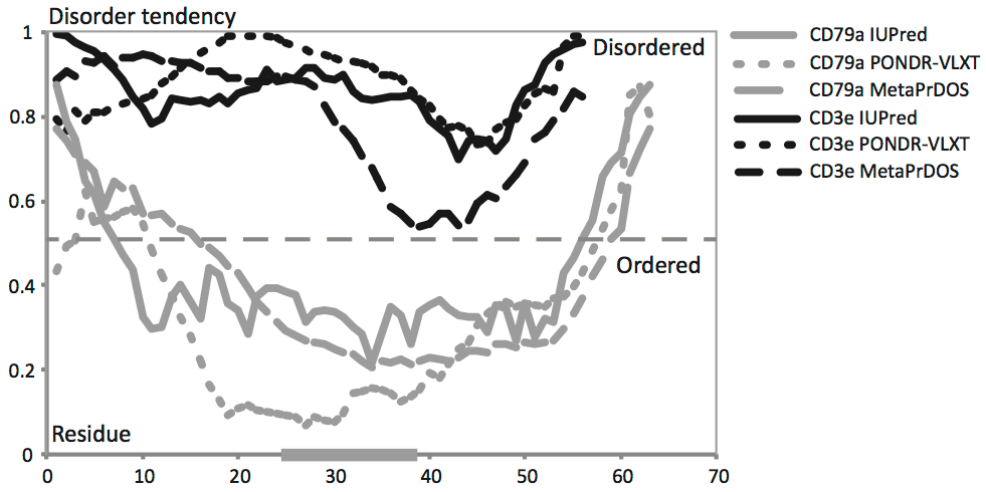


Figure 29. ***In silico*** **prediction for disorder in the cytoplasmic tails of CD79a and CD3ε using three different predictors**; IUPRed [174], PONDR-VLXT [175] and MetaPRDos [176]. ITAM in CD79a is between residue 24-37 and for CD3ε, residue 24-38, shown as a thicker line.

Biochemical/biophysical methods

IDPs do not behave biochemically and biophysically as globular proteins and some of these properties can be used to identify and investigate them. For example, they are often resistant to heat-denaturation and hence, purification can be made simple. H/D-exchange is very fast in IDPs since almost all side chains are exposed to solvent and this also makes the proteins more prone to protease activity. In many techniques based on size or hydrodynamic behaviour, IDPs behave as if they are larger than they are; Size-exclusion chromatography, SDS-poly acrylamide gel electrophoresis (SDS-PAGE), dynamic light scattering (DLS) and NMR diffusion experiments.

Long-range-interactions

NMR can examine many different properties of the proteins and is particularly well suited for IDPs. In spite of their highly dynamic nature, IDPs may exhibit transient or persistent long-range tertiary structure that may be related to biological activity or inhibit protein from proteolysis or amyloidosis (such as in Aβ or plack formation of α-synnuclein in Alzheimer's and Parkinson's disease) [58].

Long-range interactions within a protein or between proteins are often probed by the interaction between the observed spin and an unpaired electron from an artificially

introduced nitroxid group, so-called paramagnetic relaxation enhancement, PRE, also called spin-labels. In its oxidized state the label is paramagnetic and relaxation is induced via dipolar coupling to the unpaired electron spin, it bleaches out signals in a radius of about 20Å. In the reduced state, the label is diamagnetic, not active and the spectrum of this is recorded as reference and the residues on the protein close in space with the PRE the ones showing decrease in peak intensity.

Dipolar couplings between two nuclei may be strong and depends on the geometry between them but averages zero due to equal possibility in free solution [53, 177]. If the protein is dissolved in a solution of weakly aligning media, such as polyacrylamide gels, phages or lipid bicelles, the residual dipolar couplings (RDC) between two nuclei are reintroduced and may be recorded. This gives an angle between each atom pairs and is used for structure calculation in globular proteins. In IDPs the RDC values are compared to a database of random coil angles and since secondary structures are sensitive to specific backbone dihedral angle propensities for structure can be seen [53, 178]. Often different techniques are combined in computational programs such as RDC and PRE data are combined in ASTEROIDS [179] and the combination of RDC and SAXA in the making of ths structure of p53, see *figure 8*.

Nuclear overhauser-effect, NOE, is the measuring of cross relaxation through space between two nuclei (the analogous dipolar interaction between nuclear spin and electron spin in PRE). The intensity of the signal decays with distance, r^6, and from this, spacial relationships can be calculated [156]. This is the primary information used to calculate globular structures, whereas for IDPs they are rare to find, except for sequential NOEs [58].

Relaxation

Relaxation is a measurement of the intrinsic movement in proteins and relaxation in globular proteins differ a lot from that in IDPs. IDPs are highly but not uniformly flexible. With NMR the dynamic data is provided for each and one of the residues and information about different mobility in different regions of the protein may be linked to function [180]. Relaxation dispersion is another NMR relaxation-technique that has more recently been used for IDPs, visualizing transition states and low populated states [181].

SAXS and CD

Other techniques used for characterizing proteins, and IDPs in particular, are small angle X-ray scattering, SAXS and circular dichroism, CD. For globular proteins, SAXS gives a low-resolution 3D structure. For IDPs, SAXS can assess the oligomeric states of protein or

protein complexes and based on a genetic algorithm and a large datasets of randomly generated models, several subgroups of protein conformations can be calculated [182]. This gives an overall shape of the ensemble of conformers for that protein. SAXS often is combined with RDC and PRE measurements [53]. CD measurements are based on different absorptions of chiral light over a range of wavelengths and the far-UV-spectrum, 180-260nm, which originates mainly from amide-bonds. The CD data gives an estimate of the amount of secondary structure and the nature of that structure, α-helix, β-sheet, or lack of structure, random coil. CD is often used initially to verify lack of secondary structure content in IDPs and monitor changes upon binding or changes in the environment [53].

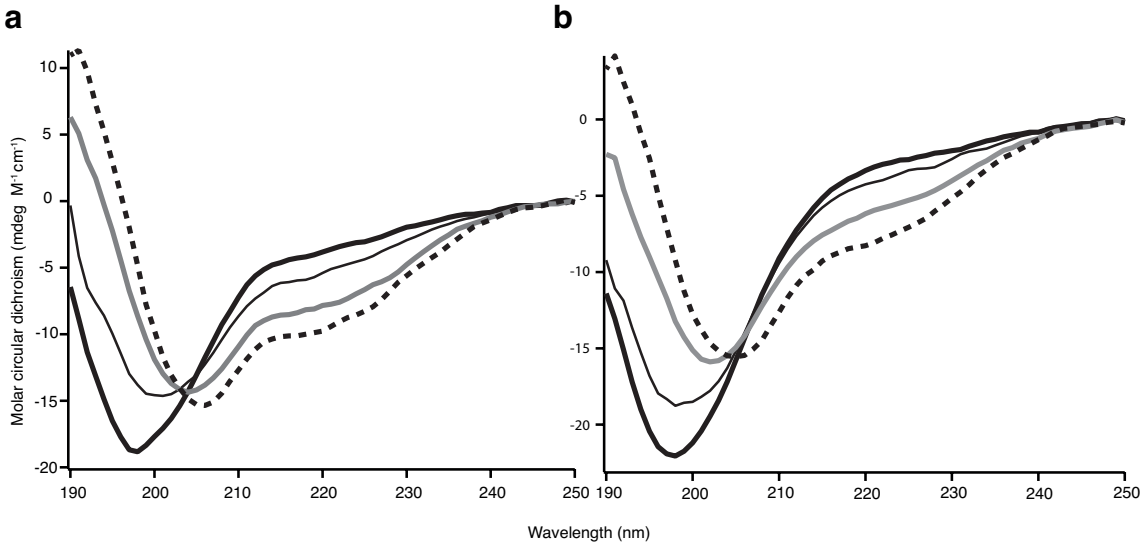


Figure 30. CD spectra for 0.38 mg/ml CD79a (a) and 0.19 mg/ml CD3ε (b) in buffer (thick line), 10 % trifluoroethanol (TFE) (thin line), 20 % TFE (grey line) and 30 % TFE (dashed line). Under native conditions both constructs showed characteristic random coil curves with a minima around 200 nm. Increasing TFE concentration gradually increased the alpha helix characteristics, with minima at 222 nm, 208 nm and a positive value at 190 nm. Picture adapted from supplementary of *paper IV*.

SUMMARY OF PUBLICATIONS:

Paper I; "The intrinsically disordered cytoplasmic domain of the T cell receptor zeta chain binds to the nef protein of simian immunodeficiency virus without a disorder-to-order transition."

In this study we show that the intrinsically disordered intracellular domain of TCR, ζ, does not undergo coupled binding and folding when interacting with the SIV Nef protein. The two proteins are known to interact from yeast 2-hybride screens [117] and SPR data [113] with two separate binding sections on ζ (SNID1, 2) to the core of Nef. The binding constant, K_d, is calculated to 4 μM [113] whereas ζ dimerization is 10 μM, [18] hence, upon complex formation, Nef breaks the (possible) dimerization of ζ. In this work, NMR experiment of free ζ show similar ^{15}NHSQC spectra as when equal amount of SIV Nef is present in the NMR tube; equal peak intensities and peak positions. Both Nef and ζ are soluble to 1mM separately, but mixed, they precipitate at concentrations > 25 μM, indicating that they affect each other and interestingly, the precipitate is mainly TCRζ.

Conclusion is that NMR data show no significant changes in chemical shift in ζ upon interaction with Nef, yet other groups have reported the interaction to be specific [117]. The nature of this binding is not understood.

Later, Walter *et al* published a crystal structure of the complex Nef_{core} – part of ζ [116] where the first ITAM in ζ adopts an α-helix-like structure in the binding cleft of Nef. This apparent contradiction points to the more complex nature of IDPs, their conformations may be very environment dependent; speculate there is a transient α-helix, which interacts with Nef. If condition is deliberately changed to select for crystals and upon crystal formation the equilibrium is shifted and eventually most ζ is trapped, in the α-helix shape, in the crystal.

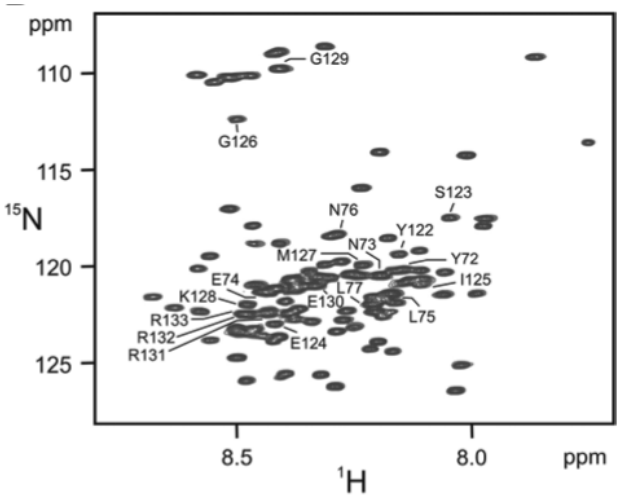


Figure 31. ^{15}N HSQC of Zeta free (grey) and Zeta-Nef 1:1 complex (black), show no visible difference in the spectra. The two spectra are overlaid and residues in the SNIS1 and SNID2 regions, know to interact with Nef, are shown. Detailed analysis of the intensities show a marginally decrease of about 10-15% in the spectrum of complex, reasoned by viscosity increase and the larger size of the complex. (Unpublished data). Figure from *paper 1*.

Paper II; "Backbone resonance assignment of Staphylococcal Enterotoxin H."

The basis for most BioNMR work is the assignment, this is the key to couple changes seen in the spectra due to a physical difference to a residue or area in the protein. Here, backbone assignment of SEH has been completed mainly from 2mg of a triple-labelled (^{2}H, ^{13}C and ^{15}N) sample. The amide protons were back-exchanged to H^{1} for visualization in the NMR experiments and this was not complete due to a very strong β-sheet. Thus, a double-labeled sample (^{13}C and ^{15}N) was used for these signals as well as for ^{15}N NOESY-HSQC experiment to verify the assignment. This is the first time a superantigen has been assigned and it opens the door to many NMR experiments such as interaction studies. This is what continued in *paper 3*.

The lack of back exchange of amide protons in the β-sheet (β strand 9, 10, 12) shows the stability of protein; 50°C shaking for 60h and the H-bonds within the β-sheet do not let go, see *figure 32*. The assignment was done using CCPN [183] and NMRPipe [184] and data deposited to the BioMagneticResonanceBank (BMRB number 16146).

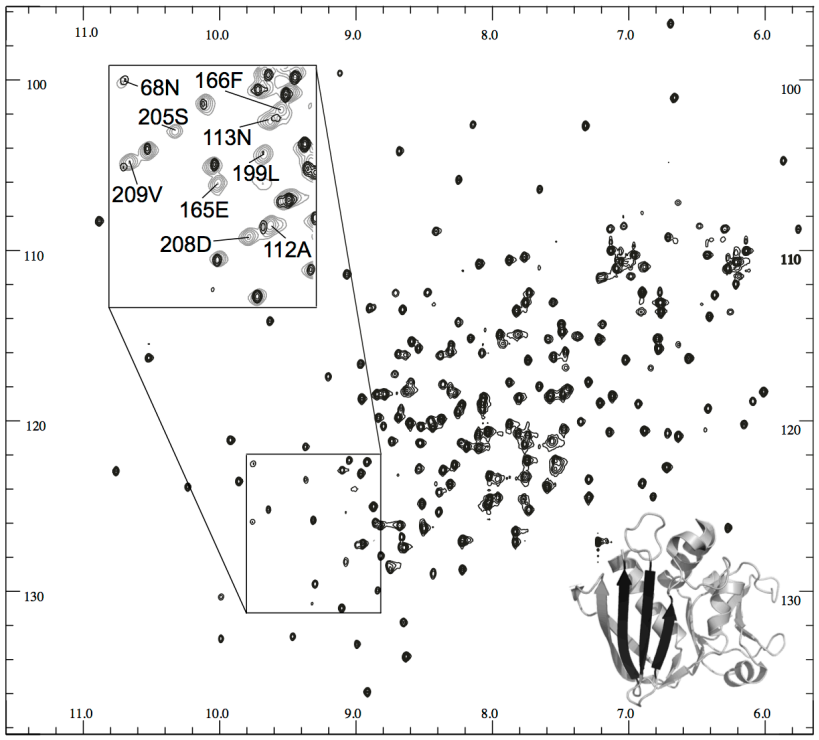


Figure 32. The effect of deuterium labelling. Line widths are much smaller giving sharper lines and better peak separation, due to the relaxation properties of deuterium. ^{15}N HSQC of SEH, black spectrum is of the deuterium (^{2}H, ^{13}C and ^{15}N) labelled sample and grey of the double labelled (^{13}C and ^{15}N). In the lower right corner the where position of the strong β -sheets on SEH is visualized in black. Picture adapted from [157].

Paper III; "The structure of superantigen complexed with TCR and MHC reveals novel insights into superantigenic T cell activation."

Here we, for the first time ever, show a three-dimensional structure of a superantigen, SEH, crosslinking both MHC class II presenting a peptide to a full length TCR. The structure is inline with previous data from cell upregulation studies [39] and SPR [136] showing the main contact between SEH and TCR is from the α-chain of TCR, something never seen for any other superantigen. The crystal structure shows a similar interaction surface on SEH as other superantigens-TCR structures, this is also confirmed by NMR data of the complex in solution. The exact amino acids crucial for SEH-TCR interaction are conserved amongst other related SAgs. Comparing the important residues for this interaction, it suggests that other superantigens possibly can up regulate cells dependent on the α-chain of TCR.

The NMR chemical shift changes in peak position and intensity were analysed to map the SEH residues the most affected by the presence of the TCR (TRAV27/TRVB19).

For chemical shifts changes, the peak positions in ^{1}H and ^{15}N dimensions were weighted together and the peak intensities were normalized and the difference between peaks in spectrum of free SEH and spectrum of SEH with addition of 21 mM TCR (1:6 SEH:TCR) were analysed see *figure 28*. The most affected residues correspond well with the interaction surface seen in the x-ray structure.

Protein production for this work was substantial; five different genes were expressed in separate cell lines. One, SEH, transcribed and translated into a very toxic protein and the other four, TCR and MHC made inclusion bodies, which had to be purified and refolded to the correct heterodimers.

Our hope is that this finding will open up new possibilities and widen the view in the field of superantigen-biology, as well as understanding protein-protein interactions.

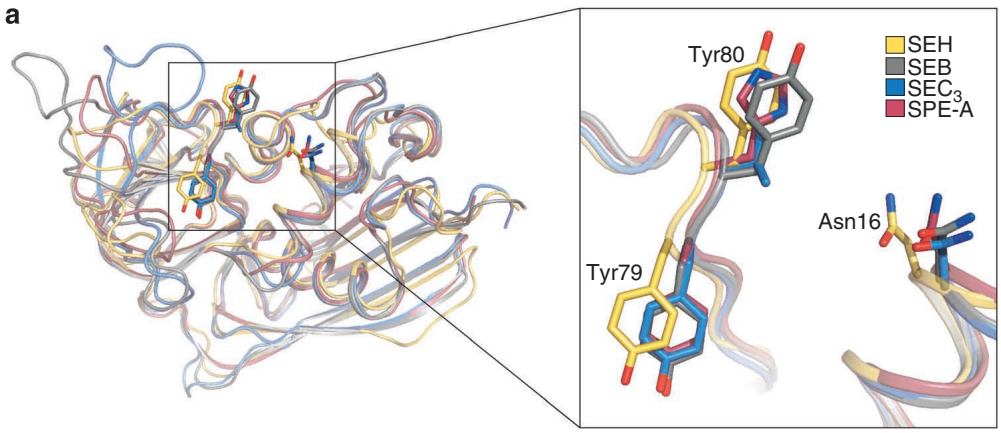


Figure 33. The three most involved residues in the SEH-TRAV27 interaction are conserved structurally within several other superantigens; maybe they also can engage the TCR through the α-chain. Picture from *paper III* [1].

Paper IV; "Highly efficient NMR assignment of intrinsically disordered proteins: application to B- and T cell receptor domains"

The field of IDPs is still in its beginning and to accumulate data and knowledge, appropriate methods are needed. IDPs require NMR assignment in many different conditions, with different mutations and additions of spin labels for a complete understanding of their nature; hence, fast protein productions and assignments are needed. In this paper we demonstrate a new methodology based on techniques proven fast and reliable, cell free protein synthesis (CFPS) and fast NMR.

The proteins are produced using CFPS where the translation machinery from *E. coli* is extracted and added to a test tube for more control over process. DNA encoding the protein together with necessary molecules is added and especially for protease-sensitive IDPs, protease inhibitors are present during protein production. IDPs suffer from low dispersion in the H-dimension circumvented by multidimensional NMR, which traditionally takes long time to record. We use non-uniform sampling, NUS, together with fast-pulsing experiments to decrease the acquisition time. Data is recorded using Targeted acquisition, TA and continuously being processed, analysed and validated, a completely new approach minimizing false assignments. We call this process TANSY Targeted Acquisition NMR SpectroscopY.

Using this approach CD79a has been assigned in one hour, about hundred times faster than conventional methods. Even at very low concentrations, 30 μM, assignment is possible. Assignment of CD79a in 20% TFE as well as 'native' condition (phosphate buffer) show relatively small but unambiguous α-helical secondary shifts, notably enhanced by the TFE. An MORF such as this α-helix in the ITAM region of CD79a has previously been reported in interaction with Lyn [185] and hence, could be the active form of the protein.

ACKNOWLEDGEMENT

It has been a long journey, and the last months I have tried to restart my brains to remember what I once knew… As I read all the reference articles and look at our data again, I realize that there is so much fun science out there. I also recognize that now when I am about to leave with this thesis in my hand, that I will miss many, many people; you are what make working the most fun. I want to say thanks to…

Vladislav, my advisor, you know NMR spectroscopy like no one else. You are a thoughtful man and a very good scientist, I have learnt a lot from you. Especially I have enjoyed all the long discussion we´ve had; some about science, about all potential reasons and scenarios why the proteins we investigate behave as they do. Some about everything else in the world; politics, military history, philosophy and the nature of human beings, you are wise.

Karin L-P, min biträdande handledare som lyste upp med ett superspännande projekt som sökte fortsättning, det blev ju mer än bara analysera ett par NMR spektra! Du är väldigt rak och tydlig både i tal och skrift, jag uppskattar det och har lärt mycket av att jobba med dig, om mig själv och hur vetenskapsvärlden fungerar. Jag önskar dig all lycka i Lund!

Göran, min biträdande handledare som hjälpte mig i SEH-projektet och alltid funnits för frågor om NMR eller allt kring doktorandlivet. Du är en hejare på att snacka skit om klimatförändringarna och så många gånger du/vi har räddat världen runt lunchbordet, varför är vi inte politiker!?

Martin & **Jan,** mina examinatorer som godkänt div. kurser jag varit med på, samt ätit en och annan kakbit ihop....

Michelle for accepting to be my opponent and the invitation to present in your lab 2010, it was a great experience.

Anders P, det har sagts förut, men är så sant, du är en klippa på allt som har med biokemi-labbande att göra och jag är oerhört glad att jag lärt känna och lärt av dig. Jag kommer sakna dina utbrott över knepiga samarbetspartners, glada utbrott när det gått bra och din omtänksamhet.

Linnéa, det har varit fantastiskt att vara siamesisk lab-tvilling med dig! Allt blir så mycket bättre när man är två, och nederlagen känns inte lika hårda. Jag tror vi kommer jobba mer ihop, allt annat känns fel.

Cissi, du är tryggheten själv, så varm och okonstlad, jag blir glad varje gång jag ser dig! Du svara lugnt på alla knepiga NMR relaterade frågor och tror jag att vi diskuterat det mesta annat över en kopp kaffe. Mer skidåkning nästa år!

Maxim, fint samarbete om IDPs, det är inspirerande att jobba med dig, du kan NMR! Hoppas jag kan få lära dig bergsklättra i vår!

Johan, vi har jobbat parallellt hela tiden, synd att vi aldrig hade något jobb ihop. Många luncher och mycket kaffe har det blivit, du kan så mycket och kommer ihåg vad du läst, jag har tappat hakan många gånger.

Joakim, så himla skoj att du fortsätter på projekten, det kommer gå strålande. Du är kunnig och ger allt den tid det behöver, väldigt klokt, du kommer komma fram till många spännande saker, sen har du en bra syn på det mest i livet.

Jonas & **Daniel** det var skönt att ha er med från Moskvas förort och marknader till båtresan St. Petersburg med tillhörande förvirring, mat innehållande kål i alla former och NMR teorier. **Diana, Weixiao**, utan er hade jag aldrig kommit igenom Spin Dynamics.

Lars-Gunnar, Patrik & **Anders** – NMR maskin-experter som håller maskinerna gående.

Maria J, Victor, Anastasia – from the old days, but I remember you as if it was yesterday…

PEST-gruppen, ett bra forum att vidga sitt vetenskapliga tänk i en skön grupp; **Karin E, Karin L-P, Urszula, Cissi, Therese, Marcus** och **Jenny. Madde** nu jäklar Madde, nu finns tid för bus och skoj! Tack för alla ärvda saker och luncher, du har en skön energi! **Karin R,** kanske det smartaste jag gjort som doktorand var att anta dig som min ex-jobbare, det var superkul att jobba med dig, du o du kan gå hur långt du vill!!

Anna F, det var himla skoj att undervisa med dig, du har organiseringen och kollen jag saknar, vi samarbetar bra. Hoppas det kan komma mer av det! Sen kanske vi båda har mer tid att ses nu när vi är klara, du har grill, vi har mat och dryck och i utby finns fina ängar…

Resten av Lundberg Lab och alla som där tillhör eller tillhört. Det har varit skoj att se var era data tar er och ni dem. **Linda**, tack för det fina samarbetet på Biokemi2, lycka till med din avhandling.

Gabriel Brandt, Nilou Ataie and **Gregory Petsko**, at Brandeis, for presenting me to the structural biology around TCR and making science so exciting.

Alla stiftelser och organisationer som under åren gett mig stipendier att avsluta dessa studier och att åka på konferenser och kurser inkluderande skidåkning i klippiga bergen, bada i sop-

vatten i Pireaus och träffa en vilsen brasilianare med gitarr på axeln i St. Petersburg samt många andra spännande forskare.

Barrie – tech support and house guarding.

Isabelle – för en vacker framsida och innan dess tillkomst, fina diskussioner om garnnystan, serpentine och rymdskeppet i bakgrunden…

Camilla – du drar ut mig till klippan och får mig bli mänsklig och leva nu.

Hanna L, Caroline, Ingrid, Sara och resten av TBi-99 gänget, jag trodde aldrig jag skulle fortsätta studera, inte utan er hjälp och avbrott med roliga äventyr. Men ni finns kvar och det är jag så himla tacksam och glad för, ni betyder massor för mig!

Hanna, Sanna, Vanja, Anton, Gustav, Torgeir, Arvid och **Peppe,** gamla vänner är äldst, det gör mig så glad att ha er. Även om vardagslivet inte spelar roll när vi ses så kanske ni nu kan förstå lite mer av vad jag pysslat med de senaste 7 åren!

Hanna C, det är tidslöst och härligt att komma hem, till er, jag blir så glad av att se dig!

Kicki, vilka goa' luncher på franska vi hade, ett skönt avbrott i labbandet. Nu kan vi ha skoj, ni är hemma och jag färdig!

US friends: Chad, Blair, Colby, Paula, Adam, Jessica, Christo, discussing high and low, science and life and being there to warm my heart.

Corinne & Angelo Cilli for hosting me in Steamboat, CO and showing me cowboy-land and real powder-skiing

Craig & Sharon for being the most wonderful parent-in-law. Especially thanks for the service we got the first three weeks I was back at work, including, food, drinks and drop-off/pick-up at dagis. You are the best!

Familj, för att ni är den trygga mur jag kan luta mig emot. För stort och smått, middagar, lån av maskiner och att ni tror på mig! **Farfar,** du frågar var gång vi talas vid "hur det går med forskningen och ryssen", jag är imponerad att du hänger med och det är härligt att ha en hejarklack hemma i Borlänge. **Mamma** och **Pappa,** ni har stöttat och hejat på mig i allt jag gjort, egentligen utan att riktigt säga det rakt ut, för er är det bara en självklarhet att jag klarar av det jag vill. Ni är de bästa föräldrar man kan ha!

Eric, Alexander & **Elliot,** att komma hem till er är alltid det bästa**,** ni är det viktigaste för mig!

REFERENCES

1. Saline, M., et al., *The structure of superantigen complexed with TCR and MHC reveals novel insights into superantigenic T cell activation.* Nat Commun, 2010. **1**: p. 119.
2. Goldsby, R.A., et al., *Kuby immunology*. 4th ed. 2000, New York: W.H. Freeman. xxv, 670 p.
3. Brändén, C.-I. and J. Tooze, *Introduction to protein structure*. 2nd ed. 1999, New York: Garland Pub. xiv, 410 p.
4. Kurosaki, T., *Regulation of B cell fates by BCR signaling components.* Curr Opin Immunol, 2002. **14**(3): p. 341-7.
5. van der Merwe, P.A. and O. Dushek, *Mechanisms for T cell receptor triggering.* Nat Rev Immunol, 2011. **11**(1): p. 47-55.
6. Engelhorn, M.E., et al., *Autoimmunity and tumor immunity induced by immune responses to mutations in self.* Nat Med, 2006. **12**(2): p. 198-206.
7. Barre-Sinoussi, F., et al., *Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS).* Science, 1983. **220**(4599): p. 868-71.
8. Fraser, J.D. and T. Proft, *The bacterial superantigen and superantigen-like proteins.* Immunol Rev, 2008. **225**: p. 226-43.
9. LeBien, T.W. and T.F. Tedder, *B lymphocytes: how they develop and function.* Blood, 2008. **112**(5): p. 1570-80.
10. Rudolph, M.G., R.L. Stanfield, and I.A. Wilson, *How TCRs bind MHCs, peptides, and coreceptors.* Annu Rev Immunol, 2006. **24**: p. 419-66.
11. Call, M.E. and K.W. Wucherpfennig, *The T cell receptor: critical role of the membrane environment in receptor assembly and function.* Annu Rev Immunol, 2005. **23**: p. 101-25.
12. Sun, Z.Y., et al., *Solution structure of the CD3epsilondelta ectodomain and comparison with CD3epsilongamma as a basis for modeling T cell receptor topology and signaling.* Proc Natl Acad Sci U S A, 2004. **101**(48): p. 16867-72.
13. Sun, Z.J., et al., *Mechanisms contributing to T cell receptor signaling and assembly revealed by the solution structure of an ectodomain fragment of the CD3 epsilon gamma heterodimer.* Cell, 2001. **105**(7): p. 913-23.
14. Ghendler, Y., et al., *Differential thymic selection outcomes stimulated by focal structural alteration in peptide/major histocompatibility complex ligands.* Proc Natl Acad Sci U S A, 1998. **95**(17): p. 10061-6.
15. Call, M.E. and J.J. Chou, *A view into the blind spot: solution NMR provides new insights into signal transduction across the lipid bilayer.* Structure, 2010. **18**(12): p. 1559-69.
16. Call, M.E., K.W. Wucherpfennig, and J.J. Chou, *The structural basis for intramembrane assembly of an activating immunoreceptor complex.* Nat Immunol, 2010. **11**(11): p. 1023-9.
17. Kuhns, M.S., et al., *Evidence for a functional sidedness to the alphabetaTCR.* Proc Natl Acad Sci U S A, 2010. **107**(11): p. 5094-9.

18. Sigalov, A., D. Aivazian, and L. Stern, *Homooligomerization of the cytoplasmic domain of the T cell receptor zeta chain and of other proteins containing the immunoreceptor tyrosine-based activation motif.* Biochemistry, 2004. **43**(7): p. 2049-61.
19. Reth, M., *Antigen receptor tail clue.* Nature, 1989. **338**(6214): p. 383-4.
20. Viret, C., F.S. Wong, and C.A. Janeway, Jr., *Designing and maintaining the mature TCR repertoire: the continuum of self-peptide:self-MHC complex recognition.* Immunity, 1999. **10**(5): p. 559-68.
21. Krogsgaard, M., et al., *Agonist/endogenous peptide-MHC heterodimers drive T cell activation and sensitivity.* Nature, 2005. **434**(7030): p. 238-43.
22. Tolar, P., et al., *The molecular assembly and organization of signaling active B-cell receptor oligomers.* Immunol Rev, 2009. **232**(1): p. 34-41.
23. Yang, J. and M. Reth, *The dissociation activation model of B cell antigen receptor triggering.* FEBS Lett, 2010. **584**(24): p. 4872-7.
24. http://www.imgt.org/.
25. Wu, T.T. and E.A. Kabat, *An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementarity.* J Exp Med, 1970. **132**(2): p. 211-50.
26. Davis, M.M., et al., *Ligand recognition by alpha beta T cell receptors.* Annu Rev Immunol, 1998. **16**: p. 523-44.
27. Bjorkman, P.J., et al., *Structure of the human class I histocompatibility antigen, HLA-A2.* Nature, 1987. **329**(6139): p. 506-12.
28. Jardetzky, T.S., et al., *Three-dimensional structure of a human class II histocompatibility molecule complexed with superantigen.* Nature, 1994. **368**(6473): p. 711-8.
29. Stadinski, B.D., et al., *A role for differential variable gene pairing in creating T cell receptors specific for unique major histocompatibility ligands.* Immunity, 2011. **35**(5): p. 694-704.
30. Hare, B.J., et al., *Structure, specificity and CDR mobility of a class II restricted single-chain T-cell receptor.* Nat Struct Biol, 1999. **6**(6): p. 574-81.
31. Petersson, K., et al., *Crystal structure of a superantigen bound to MHC class II displays zinc and peptide dependence.* EMBO J, 2001. **20**(13): p. 3306-12.
32. Otwinowski, Z., et al., *Crystal structure of trp repressor/operator complex at atomic resolution.* Nature, 1988. **335**(6188): p. 321-9.
33. Li, H., et al., *The structural basis of T cell activation by superantigens.* Annu Rev Immunol, 1999. **17**: p. 435-66.
34. Xu, S.X. and J.K. McCormick, *Staphylococcal superantigens in colonization and disease.* Front Cell Infect Microbiol, 2012. **2**: p. 52.
35. Qu, H.Q., et al., *Ancestry informative marker set for han chinese population.* G3 (Bethesda), 2012. **2**(3): p. 339-41.
36. Ferry, T., et al., *Early diagnosis of staphylococcal toxic shock syndrome by detection of the TSST-1 Vbeta signature in peripheral blood of a 12-year-old boy.* Pediatr Infect Dis J, 2008. **27**(3): p. 274-7.
37. Swaminathan, S., et al., *Crystal structure of staphylococcal enterotoxin B, a superantigen.* Nature, 1992. **359**(6398): p. 801-6.

38. Nilsson, H., et al., *Staphylococcal enterotoxin H displays unique MHC class II-binding properties.* J Immunol, 1999. **163**(12): p. 6686-93.
39. Petersson, K., et al., *Staphylococcal enterotoxin H induces V alpha-specific expansion of T cells.* J Immunol, 2003. **170**(8): p. 4148-54.
40. Fraser, J.D., *Clarifying the mechanism of superantigen toxicity.* PLoS Biol, 2011. **9**(9): p. e1001145.
41. Borghaei, H., et al., *Phase I dose escalation, pharmacokinetic and pharmacodynamic study of naptumomab estafenatox alone in patients with advanced cancer and with docetaxel in patients with advanced non-small-cell lung cancer.* J Clin Oncol, 2009. **27**(25): p. 4116-23.
42. Sigler, P.B., *Transcriptional activation. Acid blobs and negative noodles.* Nature, 1988. **333**(6170): p. 210-2.
43. Oldfield, C.J., et al., *Comparing and combining predictors of mostly disordered proteins.* Biochemistry, 2005. **44**(6): p. 1989-2000.
44. Ward, J.J., et al., *Prediction and functional analysis of native disorder in proteins from the three kingdoms of life.* J Mol Biol, 2004. **337**(3): p. 635-45.
45. Iakoucheva, L.M., et al., *Intrinsic disorder in cell-signaling and cancer-associated proteins.* J Mol Biol, 2002. **323**(3): p. 573-84.
46. Uversky, V.N. and A.K. Dunker, *Understanding protein non-folding.* Biochim Biophys Acta, 2010. **1804**(6): p. 1231-64.
47. Tompa, P. and M. Fuxreiter, *Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions.* Trends Biochem Sci, 2008. **33**(1): p. 2-8.
48. Mittag, T., L.E. Kay, and J.D. Forman-Kay, *Protein dynamics and conformational disorder in molecular recognition.* J Mol Recognit, 2010. **23**(2): p. 105-16.
49. Sigalov, A.B., A.V. Zhuravleva, and V.Y. Orekhov, *Binding of intrinsically disordered proteins is not necessarily accompanied by a structural transition to a folded form.* Biochimie, 2007. **89**(3): p. 419-21.
50. Cheng, Y., et al., *Mining alpha-helix-forming molecular recognition features with cross species sequence alignments.* Biochemistry, 2007. **46**(47): p. 13468-77.
51. Preusser, A., L. Briese, and D. Willbold, *Presence of a helix in human CD4 cytoplasmic domain promotes binding to HIV-1 Nef protein.* Biochem Biophys Res Commun, 2002. **292**(3): p. 734-40.
52. Norholm, A.B., et al., *The intracellular distal tail of the Na+/H+ exchanger NHE1 is intrinsically disordered: implications for NHE1 trafficking.* Biochemistry, 2011. **50**(17): p. 3469-80.
53. Peter, T., *Structure and function of disordered proteins.* Taylor & Francis group, 2009.
54. Ramelot, T.A., L.N. Gentile, and L.K. Nicholson, *Transient structure of the amyloid precursor protein cytoplasmic tail indicates preordering of structure for binding to cytosolic factors.* Biochemistry, 2000. **39**(10): p. 2714-25.
55. Dyson, H.J. and P.E. Wright, *Elucidation of the protein folding landscape by NMR.* Methods Enzymol, 2005. **394**: p. 299-321.
56. Sugase, K., H.J. Dyson, and P.E. Wright, *Mechanism of coupled folding and binding of an intrinsically disordered protein.* Nature, 2007. **447**(7147): p. 1021-5.

57. Mittag, T., et al., *Dynamic equilibrium engagement of a polyvalent ligand with a single-site receptor.* Proc Natl Acad Sci U S A, 2008. **105**(46): p. 17772-7.
58. Dyson, H.J., *Expanding the proteome: disordered and alternatively folded proteins.* Q Rev Biophys, 2011. **44**(4): p. 467-518.
59. Tompa, P., C. Szasz, and L. Buday, *Structural disorder throws new light on moonlighting.* Trends Biochem Sci, 2005. **30**(9): p. 484-9.
60. Dyson, H.J. and P.E. Wright, *Intrinsically unstructured proteins and their functions.* Nat Rev Mol Cell Biol, 2005. **6**(3): p. 197-208.
61. Hegyi, H. and P. Tompa, *Intrinsically disordered proteins display no preference for chaperone binding in vivo.* PLoS Comput Biol, 2008. **4**(3): p. e1000017.
62. Gsponer, J., et al., *Tight regulation of unstructured proteins: from transcript synthesis to protein degradation.* Science, 2008. **322**(5906): p. 1365-8.
63. Babu, M.M., et al., *Intrinsically disordered proteins: regulation and disease.* Curr Opin Struct Biol, 2011. **21**(3): p. 432-40.
64. Metallo, S.J., *Intrinsically disordered proteins are potential drug targets.* Curr Opin Chem Biol, 2010. **14**(4): p. 481-8.
65. Uversky, V.N., C.J. Oldfield, and A.K. Dunker, *Intrinsically disordered proteins in human diseases: introducing the D2 concept.* Annu Rev Biophys, 2008. **37**: p. 215-46.
66. Wells, M., et al., *Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain.* Proc Natl Acad Sci U S A, 2008. **105**(15): p. 5762-7.
67. Dong, Q., et al., *Analysis and prediction of protein local structure based on structure alphabets.* Proteins, 2008. **72**(1): p. 163-72.
68. Sigalov, A.B., et al., *The intrinsically disordered cytoplasmic domain of the T cell receptor zeta chain binds to the nef protein of simian immunodeficiency virus without a disorder-to-order transition.* Biochemistry, 2008. **47**(49): p. 12942-4.
69. Uversky, V.N., J.R. Gillespie, and A.L. Fink, *Why are "natively unfolded" proteins unstructured under physiologic conditions?* Proteins, 2000. **41**(3): p. 415-27.
70. Engels, N., B. Wollscheid, and J. Wienands, *Association of SLP-65/BLNK with the B cell antigen receptor through a non-ITAM tyrosine of Ig-alpha.* Eur J Immunol, 2001. **31**(7): p. 2126-34.
71. Kraus, M., et al., *Interference with immunoglobulin (Ig)alpha immunoreceptor tyrosine-based activation motif (ITAM) phosphorylation modulates or blocks B cell development, depending on the availability of an Igbeta cytoplasmic tail.* J Exp Med, 2001. **194**(4): p. 455-69.
72. Pelanda, R., et al., *B cell progenitors are arrested in maturation but have intact VDJ recombination in the absence of Ig-alpha and Ig-beta.* J Immunol, 2002. **169**(2): p. 865-72.
73. Conley, M.E., et al., *Primary B cell immunodeficiencies: comparisons and contrasts.* Annu Rev Immunol, 2009. **27**: p. 199-227.
74. Davis, R.E., et al., *Chronic active B-cell-receptor signalling in diffuse large B-cell lymphoma.* Nature, 2010. **463**(7277): p. 88-92.
75. Call, M.E. and K.W. Wucherpfennig, *Common themes in the assembly and architecture of activating immune receptors.* Nat Rev Immunol, 2007. **7**(11): p. 841-50.

76. Monroe, J.G., *ITAM-mediated tonic signalling through pre-BCR and BCR complexes.* Nat Rev Immunol, 2006. **6**(4): p. 283-94.
77. Batista, F.D., D. Iber, and M.S. Neuberger, *B cells acquire antigen from target cells after synapse formation.* Nature, 2001. **411**(6836): p. 489-94.
78. James, J.R. and R.D. Vale, *Biophysical mechanism of T-cell receptor triggering in a reconstituted system.* Nature, 2012. **487**(7405): p. 64-9.
79. Dushek, O., et al., *Antigen potency and maximal efficacy reveal a mechanism of efficient T cell activation.* Sci Signal, 2011. **4**(176): p. ra39.
80. Zola, H., et al., *CD molecules 2006--human cell differentiation molecules.* J Immunol Methods, 2007. **319**(1-2): p. 1-5.
81. Edidin, M., *Lipids on the frontier: a century of cell-membrane bilayers.* Nat Rev Mol Cell Biol, 2003. **4**(5): p. 414-8.
82. Brown, D.A. and J.K. Rose, *Sorting of GPI-anchored proteins to glycolipid-enriched membrane subdomains during transport to the apical cell surface.* Cell, 1992. **68**(3): p. 533-44.
83. Zech, T., et al., *Accumulation of raft lipids in T-cell plasma membrane domains engaged in TCR signalling.* EMBO J, 2009. **28**(5): p. 466-76.
84. Horejsi, V., *Lipid rafts and their roles in T-cell activation.* Microbes Infect, 2005. **7**(2): p. 310-6.
85. Kung, C., *A possible unifying principle for mechanosensation.* Nature, 2005. **436**(7051): p. 647-54.
86. Tolar, P., H.W. Sohn, and S.K. Pierce, *The initiation of antigen-induced B cell antigen receptor signaling viewed in living cells by fluorescence resonance energy transfer.* Nat Immunol, 2005. **6**(11): p. 1168-76.
87. Fischer, K., et al., *Antigen recognition induces phosphatidylserine exposure on the cell surface of human CD8+ T cells.* Blood, 2006. **108**(13): p. 4094-101.
88. Cochran, J.R., T.O. Cameron, and L.J. Stern, *The relationship of MHC-peptide binding and T cell activation probed using chemically defined MHC class II oligomers.* Immunity, 2000. **12**(3): p. 241-50.
89. Irvine, D.J., et al., *Direct observation of ligand recognition by T cells.* Nature, 2002. **419**(6909): p. 845-9.
90. Sigalov, A.B., *The SCHOOL of nature: I. Transmembrane signaling.* Self Nonself, 2010. **1**(1): p. 4-39.
91. Lillemeier, B.F., et al., *TCR and Lat are expressed on separate protein islands on T cell membranes and concatenate during activation.* Nat Immunol, 2010. **11**(1): p. 90-6.
92. Yokosuka, T., et al., *Newly generated T cell receptor microclusters initiate and sustain T cell activation by recruitment of Zap70 and SLP-76.* Nat Immunol, 2005. **6**(12): p. 1253-62.
93. Varma, R., et al., *T cell receptor-proximal signals are sustained in peripheral microclusters and terminated in the central supramolecular activation cluster.* Immunity, 2006. **25**(1): p. 117-27.
94. Yang, J. and M. Reth, *Oligomeric organization of the B-cell antigen receptor on resting cells.* Nature, 2010. **467**(7314): p. 465-9.

95. Tolar, P., et al., *The constant region of the membrane immunoglobulin mediates B cell-receptor clustering and signaling in response to membrane antigens.* Immunity, 2009. **30**(1): p. 44-55.
96. Kim, P.W., et al., *A zinc clasp structure tethers Lck to T cell coreceptors CD4 and CD8.* Science, 2003. **301**(5640): p. 1725-8.
97. Beddoe, T., et al., *Antigen ligation triggers a conformational change within the constant domain of the alphabeta T cell receptor.* Immunity, 2009. **30**(6): p. 777-88.
98. Kjer-Nielsen, L., et al., *A structural basis for the selection of dominant alphabeta T cell receptors in antiviral immunity.* Immunity, 2003. **18**(1): p. 53-64.
99. Hawse, W.F., et al., *Cutting edge: Evidence for a dynamically driven T cell signaling mechanism.* J Immunol, 2012. **188**(12): p. 5819-23.
100. Baker, B.M., et al., *Structural and dynamic control of T-cell receptor specificity, cross-reactivity, and binding mechanism.* Immunol Rev, 2012. **250**(1): p. 10-31.
101. Kuhns, M.S., M.M. Davis, and K.C. Garcia, *Deconstructing the form and function of the TCR/CD3 complex.* Immunity, 2006. **24**(2): p. 133-9.
102. Kuhns, M.S. and M.M. Davis, *The safety on the TCR trigger.* Cell, 2008. **135**(4): p. 594-6.
103. Xu, C., et al., *Regulation of T cell receptor activation by dynamic membrane binding of the CD3epsilon cytoplasmic tyrosine-based motif.* Cell, 2008. **135**(4): p. 702-13.
104. Duchardt, E., et al., *Structure induction of the T-cell receptor zeta-chain upon lipid binding investigated by NMR spectroscopy.* Chembiochem, 2007. **8**(7): p. 820-7.
105. Aivazian, D. and L.J. Stern, *Phosphorylation of T cell receptor zeta is regulated by a lipid dependent folding transition.* Nat Struct Biol, 2000. **7**(11): p. 1023-6.
106. Zhang, H., et al., *Basic residues in the T-cell receptor zeta cytoplasmic domain mediate membrane association and modulate signaling.* Proc Natl Acad Sci U S A, 2011. **108**(48): p. 19323-8.
107. Gil, D., et al., *Recruitment of Nck by CD3 epsilon reveals a ligand-induced conformational change essential for T cell receptor signaling and synapse formation.* Cell, 2002. **109**(7): p. 901-12.
108. Shi, X.S., et al., *Ca2+ regulates T-cell receptor activation by modulating the charge property of lipids.* Nature, 2013. **493**(7430): p. 111-+.
109. Huse, M., et al., *Spatial and temporal dynamics of T cell receptor signaling with a photoactivatable agonist.* Immunity, 2007. **27**(1): p. 76-88.
110. Huppa, J.B., et al., *Continuous T cell receptor signaling required for synapse maintenance and full effector potential.* Nat Immunol, 2003. **4**(8): p. 749-55.
111. Ziegenfuss, J.S., et al., *Draper-dependent glial phagocytic activity is mediated by Src and Syk family kinase signalling.* Nature, 2008. **453**(7197): p. 935-9.
112. Gallo, R.C., et al., *Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS).* Science, 1983. **220**(4599): p. 865-7.
113. Kim, W.M., *Characterization of the Nef-TCR Zeta inetraction and its role in the modulation of Src Family kinaseactivity; a dissertation.* University of Massachusetts Medical School. GSBS Dissertations and thesis, 2009. **Paper 434**.
114. Kestler, H.W., 3rd, et al., *Importance of the nef gene for maintenance of high virus loads and for development of AIDS.* Cell, 1991. **65**(4): p. 651-62.

115. Geyer, M., O.T. Fackler, and B.M. Peterlin, *Structure--function relationships in HIV-1 Nef.* EMBO Rep, 2001. **2**(7): p. 580-5.
116. Kim, W.M., A.B. Sigalov, and L.J. Stern, *Pseudo-merohedral twinning and noncrystallographic symmetry in orthorhombic crystals of SIVmac239 Nef core domain bound to different-length TCRzeta fragments.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 163-75.
117. Schaefer, T.M., et al., *The T-cell receptor zeta chain contains two homologous domains with which simian immunodeficiency virus Nef interacts and mediates down-modulation.* J Virol, 2000. **74**(7): p. 3273-83.
118. Swigut, T., M. Greenberg, and J. Skowronski, *Cooperative interactions of simian immunodeficiency virus Nef, AP-2, and CD3-zeta mediate the selective induction of T-cell receptor-CD3 endocytosis.* J Virol, 2003. **77**(14): p. 8116-26.
119. Bell, I., et al., *Association of simian immunodeficiency virus Nef with the T-cell receptor (TCR) zeta chain leads to TCR down-modulation.* J Gen Virol, 1998. **79 (Pt 11)**: p. 2717-27.
120. Hiipakka, M., et al., *Inhibition of cellular functions of HIV-1 Nef by artificial SH3 domains.* Virology, 2001. **286**(1): p. 152-9.
121. Kasper, K.J., et al., *Molecular requirements for MHC class II alpha-chain engagement and allelic discrimination by the bacterial superantigen streptococcal pyrogenic exotoxin C.* J Immunol, 2008. **181**(5): p. 3384-92.
122. Petersson, K., G. Forsberg, and B. Walse, *Interplay between superantigens and immunoreceptors.* Scand J Immunol, 2004. **59**(4): p. 345-55.
123. Petersson, K., et al., *Crystal structure of a SEA variant in complex with MHC class II reveals the ability of SEA to crosslink MHC molecules.* Structure, 2002. **10**(12): p. 1619-26.
124. Leder, L., et al., *A mutational analysis of the binding of staphylococcal enterotoxins B and C3 to the T cell receptor beta chain and major histocompatibility complex class II.* J Exp Med, 1998. **187**(6): p. 823-33.
125. Cho, S., et al., *Structural basis of affinity maturation and intramolecular cooperativity in a protein-protein interaction.* Structure, 2005. **13**(12): p. 1775-87.
126. Nur-ur Rahman, A.K., et al., *The T cell receptor beta-chain second complementarity determining region loop (CDR2beta governs T cell activation and Vbeta specificity by bacterial superantigens.* J Biol Chem, 2011. **286**(6): p. 4871-81.
127. Moza, B., et al., *Structural basis of T-cell specificity and activation by the bacterial superantigen TSST-1.* EMBO J, 2007. **26**(4): p. 1187-97.
128. Li, H., et al., *Three-dimensional structure of the complex between a T cell receptor beta chain and the superantigen staphylococcal enterotoxin B.* Immunity, 1998. **9**(6): p. 807-16.
129. Fields, B.A., et al., *Crystal structure of a T-cell receptor beta-chain complexed with a superantigen.* Nature, 1996. **384**(6605): p. 188-92.
130. Fernandez, M.M., et al., *Crystal structure of staphylococcal enterotoxin G (SEG) in complex with a mouse T-cell receptor {beta} chain.* J Biol Chem, 2011. **286**(2): p. 1189-95.
131. Sundberg, E.J., et al., *Structures of two streptococcal superantigens bound to TCR beta chains reveal diversity in the architecture of T cell signaling complexes.* Structure, 2002. **10**(5): p. 687-99.

132. Gunther, S., et al., *A novel loop domain in superantigens extends their T cell receptor recognition site.* J Mol Biol, 2007. **371**(1): p. 210-21.
133. Kim, J., et al., *Toxic shock syndrome toxin-1 complexed with a class II major histocompatibility molecule HLA-DR1.* Science, 1994. **266**(5192): p. 1870-4.
134. Brosnahan, A.J. and P.M. Schlievert, *Gram-positive bacterial superantigen outside-in signaling causes toxic shock syndrome.* FEBS J, 2011. **278**(23): p. 4649-67.
135. Andersen, P.S., et al., *Role of the T cell receptor alpha chain in stabilizing TCR-superantigen-MHC class II complexes.* Immunity, 1999. **10**(4): p. 473-83.
136. Pumphrey, N., et al., *Cutting edge: Evidence of direct TCR alpha-chain interaction with superantigen.* J Immunol, 2007. **179**(5): p. 2700-4.
137. Wang, L., et al., *Crystal structure of a complete ternary complex of TCR, superantigen and peptide-MHC.* Nat Struct Mol Biol, 2007. **14**(2): p. 169-71.
138. Webb, S.R. and J. Sprent, *Response of mature unprimed CD8+ T cells to Mlsa determinants.* J Exp Med, 1990. **171**(3): p. 953-8.
139. Bueno, C., et al., *Bacterial superantigens bypass Lck-dependent T cell receptor signaling by activating a Galpha11-dependent, PLC-beta-mediated pathway.* Immunity, 2006. **25**(1): p. 67-78.
140. Purcell, E.M., Torrey, H. C., Pound, R. V, *Resonance absorbtion by nuclear magnetic moments in solids.* Physical Review, 1946(69): p. 37-38.
141. Bloch, F., Hansen, W. W., Packard, M,, *Nuclear Induction.* Physical Review, 1946(69): p. 127.
142. Williamson, M.P., T.F. Havel, and K. Wuthrich, *Solution conformation of proteinase inhibitor IIA from bull seminal plasma by 1H nuclear magnetic resonance and distance geometry.* J Mol Biol, 1985. **182**(2): p. 295-315.
143. Stern, L.J. and D.C. Wiley, *The human class II MHC protein HLA-DR1 assembles as empty alpha beta heterodimers in the absence of antigenic peptide.* Cell, 1992. **68**(3): p. 465-77.
144. Boulter, J.M., et al., *Stable, soluble T-cell receptor molecules for crystallization and therapeutics.* Protein Eng, 2003. **16**(9): p. 707-11.
145. Tsumoto, K., et al., *Practical considerations in refolding proteins from inclusion bodies.* Protein Expr Purif, 2003. **28**(1): p. 1-8.
146. Frayser, M., et al., *Empty and peptide-loaded class II major histocompatibility complex proteins produced by expression in Escherichia coli and folding in vitro.* Protein Expr Purif, 1999. **15**(1): p. 105-14.
147. Pedersen, A., et al., *Rational improvement of cell-free protein synthesis.* N Biotechnol. **28**(3): p. 218-24.
148. Carlson, E.D., et al., *Cell-free protein synthesis: applications come of age.* Biotechnol Adv, 2012. **30**(5): p. 1185-94.
149. Isaksson, L., et al., *Expression screening of membrane proteins with cell-free protein synthesis.* Protein Expr Purif, 2012. **82**(1): p. 218-25.
150. Pedersen, A., et al., *Rational improvement of cell-free protein synthesis.* N Biotechnol, 2011. **28**(3): p. 218-24.

151. Kurotani, A., et al., *Comprehensive bioinformatics analysis of cell-free protein synthesis: identification of multiple protein properties that correlate with successful expression.* FASEB J, 2010. **24**(4): p. 1095-104.
152. Staunton, D., et al., *Cell-free expression and selective isotope labelling in protein NMR.* Magn Reson Chem, 2006. **44 Spec No**: p. S2-9.
153. Petsko, G.A. and D. Ringe, *Protein structure and function*. Primers in biology. 2004, London, Sunderland, MA, Oxford: New Science Press ; Sinauer Associates ; Blackwell Pub. xxii, 195 p.
154. Rhodes, G., *Crystallography made crystal clear : a guide for users of macromolecular models*. 3rd ed. Complementary science series. 2006, Amsterdam ; Boston: Elsevier/Academic Press. xxv, 306 p.
155. Teng, Q., *Handbook of structural biology : practical NMR applications*. 2005, New York: Kluwer Academic/Plenum Publishers.
156. Levitt, M.H., *Spin dynamics : basics of nuclear magnetic resonance*. 2001, Chichester ; New York: John Wiley & Sons. xxiv, 686 p.
157. Saline, M., et al., *Backbone resonance assignment of Staphylococcal Enterotoxin H.* Biomol NMR Assign, 2009. **4**(1): p. 1-4.
158. Ito, Y. and P. Selenko, *Cellular structural biology.* Curr Opin Struct Biol, 2010. **20**(5): p. 640-8.
159. Moseley, H.N., G. Sahota, and G.T. Montelione, *Assignment validation software suite for the evaluation and presentation of protein resonance assignment data.* J Biomol NMR, 2004. **28**(4): p. 341-55.
160. Billeter, M. and V. Orekhov, *Preface: Fast NMR methods are here to stay.* Top Curr Chem, 2012. **316**: p. ix-xiv.
161. Kazimierczuk, K., et al., *Random sampling in multidimensional NMR spectroscopy.* Prog Nucl Magn Reson Spectrosc, 2010. **57**(4): p. 420-34.
162. Favier, A. and B. Brutscher, *Recovering lost magnetization: polarization enhancement in biomolecular NMR.* J Biomol NMR, 2011. **49**(1): p. 9-15.
163. Orekhov, V.Y. and V.A. Jaravine, *Analysis of non-uniformly sampled spectra with multi-dimensional decomposition.* Prog Nucl Magn Reson Spectrosc, 2011. **59**(3): p. 271-92.
164. Jaravine, V.A. and V.Y. Orekhov, *Targeted acquisition for real-time NMR spectroscopy.* J Am Chem Soc, 2006. **128**(41): p. 13421-6.
165. Zhang, H., S. Neal, and D.S. Wishart, *RefDB: a database of uniformly referenced protein chemical shifts.* J Biomol NMR, 2003. **25**(3): p. 173-95.
166. Kjaergaard, M., S. Brander, and F.M. Poulsen, *Random coil chemical shift for intrinsically disordered proteins: effects of temperature and pH.* J Biomol NMR, 2011. **49**(2): p. 139-49.
167. Bennion, B.J. and V. Daggett, *The molecular basis for the chemical denaturation of proteins by urea.* Proc Natl Acad Sci U S A, 2003. **100**(9): p. 5142-7.
168. Gast, K., et al., *Trifluoroethanol-induced conformational transitions of proteins: insights gained from the differences between alpha-lactalbumin and ribonuclease A.* Protein Sci, 1999. **8**(3): p. 625-34.

169. Myers, J.K., C.N. Pace, and J.M. Scholtz, *Trifluoroethanol effects on helix propensity and electrostatic interactions in the helical peptide from ribonuclease T1.* Protein Sci, 1998. **7**(2): p. 383-8.
170. Modig, K., et al., *Detection of initiation sites in protein folding of the four helix bundle ACBP by chemical shift analysis.* FEBS letters, 2007. **581**(25): p. 4965-4971.
171. Kjaergaard, M. and F.M. Poulsen, *Disordered proteins studied by chemical shifts.* Progress in Nuclear Magnetic Resonance Spectroscopy, 2012. **60**: p. 42-51.
172. Mittermaier, A.K. and L.E. Kay, *Observing biological dynamics at atomic resolution using NMR.* Trends Biochem Sci, 2009. **34**(12): p. 601-11.
173. Zhuravleva, A., et al., *Propagation of dynamic changes in barnase upon binding of barstar: an NMR and computational study.* J Mol Biol, 2007. **367**(4): p. 1079-92.
174. Dosztanyi, Z., et al., *IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content.* Bioinformatics, 2005. **21**(16): p. 3433-4.
175. Romero, P., et al., *Sequence complexity of disordered protein.* Proteins, 2001. **42**(1): p. 38-48.
176. Ishida, T. and K. Kinoshita, *Prediction of disordered regions in proteins based on the meta approach.* Bioinformatics, 2008. **24**(11): p. 1344-8.
177. Salmon, L., et al., *NMR characterization of long-range order in intrinsically disordered proteins.* J Am Chem Soc, 2010. **132**(24): p. 8407-18.
178. Jensen, M.R., et al., *Quantitative determination of the conformational properties of partially folded and intrinsically disordered proteins using NMR dipolar couplings.* Structure, 2009. **17**(9): p. 1169-85.
179. Nodet, G., et al., *Quantitative description of backbone conformational sampling of unfolded proteins at amino acid resolution from NMR residual dipolar couplings.* J Am Chem Soc, 2009. **131**(49): p. 17908-18.
180. Eliezer, D., *Biophysical characterization of intrinsically disordered proteins.* Curr Opin Struct Biol, 2009. **19**(1): p. 23-30.
181. Mittermaier, A. and L.E. Kay, *New tools provide new insights in NMR studies of protein dynamics.* Science, 2006. **312**(5771): p. 224-8.
182. Bernado, P., et al., *Structural characterization of flexible proteins using small-angle X-ray scattering.* J Am Chem Soc, 2007. **129**(17): p. 5656-64.
183. Vranken, W.F., et al., *The CCPN data model for NMR spectroscopy: development of a software pipeline.* Proteins, 2005. **59**(4): p. 687-96.
184. Delaglio, F., et al., *NMRPipe: a multidimensional spectral processing system based on UNIX pipes.* J Biomol NMR, 1995. **6**(3): p. 277-93.
185. Gaul, B.S., et al., *Substrate recognition by the Lyn protein-tyrosine kinase. NMR structure of the immunoreceptor tyrosine-based activation motif signaling region of the B cell antigen receptor.* J Biol Chem, 2000. **275**(21): p. 16174-82.