UNIVERSITY OF BIRMINGHAM

Project 1:

INVESTIGATING THE ROLE OF LIGAND FLEXIBILITY IN THE BINDING OF MYRISTATE TO I-FABP

and

Project 2:

STUDYING THE EFFECT OF THE BACTERIAL NDK PROTEINS ON HAEMOPOIETIC STEM CELLS

by

Inara Liepina

A combined research thesis submitted to the University of Birmingham as part of the requirement for the degree of MASTER OF RESEARCH in Molecular and Cellular Biology.

College of Life and Environmental Sciences
School of Biosciences
University of Birmingham
September 2013

SUMMARY

This is a combined thesis submitted for MRes in Molecular and Cellular Biology.

The first project is based on investigating the role of ligand flexibility in the binding of myristate to I-FABP. Myristate and palmitate bind the protein I-FABP with affinities orders of magnitude higher than similar sized less flexible ligands. This seems to contradict the medicinal chemistry textbooks where making a ligand less flexible is an accepted method for increasing binding affinity. Here computer simulations were used to investigate myristate flexibility and its free energy of binding to I-FABP, to see what role flexibility has in the recognition of myristate by I-FABP.

The second project is based on studying the effect of the bacterial NDK proteins on haemapoietic stem cells. Leukaemia patients lack immunity; thus bacterial infections worsen leukaemia prognosis. It might be that bacterial invasion can affect leukaemia cells, by rendering the latter more aggressive, or even by stimulating cell proliferation. Bacterial and eukaryotic NDK proteins are highly conserved indicating that they may also be functionally similar. Previously it was shown that additional rNM23-H1 (human NDK protein) indirectly promotes leukaemia and healthy stem cell survival and proliferation. Here it was investigated if bacterial NDK proteins can show a similar effect.

ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my supervisors Dr. Peter Winn and Dr. Farhat Khanim for the continuous support of my MRes study and research.

My sincere thanks also go to Prof. Chris Bunce, my fellow lab mates Laura Cronin, Katarzyna Koczula, Andy Southam and Nikos Batis.

I would like to thank Miles Rowley and Dr. Kay Van Damme for investing time and helping me to improve my English language throughout my thesis and to explain my mistakes. This thesis was copy edited for conventions of language, spelling and grammar by Miles Rowley and Dr. Kay van Damme.

I thank my fellow course mates, especially Jelena Sostare and Monika Kopf.

Last but not the least; I would like to thank my boyfriend Kamil Trykowski for his patience and support through my studies.

Project 1:

INVESTIGATING THE ROLE OF LIGAND FLEXIBILITY IN THE BINDING OF MYRISTATE TO I-FABP

ABSTRACT

Myristate and palmitate bind the protein I-FABP with affinities orders of magnitude higher than similarly sized less flexible ligands. This seems to contradict the medicinal chemistry textbooks where making a ligand less flexible is an accepted method for increasing binding affinity. Here computer simulations were used to investigate myristate flexibility and its free energy of binding to I-FABP, to see what role flexibility has in the recognition of myristate by I-FABP.

The difference in the binding free energy between palmitate and myristate on binding to I-FABP was calculated using thermodynamic integration. The calculated value of -8.8 kJ/mol differs from the experimental value of approximately -1.5 kJ/mol. It is unclear whether this discrepancy represents a deficiency in the force field; too short a simulation time or indeed a problem with the experiment. However, the rough qualitative agreement suggests that it is still appropriate to use the system as a model to investigate whether flexibility is indeed detrimental for the binding of a myristate-type molecule to I-FABP.

Further simulations show that, for most of the dihedral angles of myristate, there is a minimal change between the I-FABP bound and solution forms. However, there is a high increase in rigidity and changes in conformational preference, in the head group of myristate. An increase of the flexibility at 6^{th} dihedral is also seen. These results will form the basis for investigating the effect of conformationally restricting myristate on its binding to I-FABP *in silico*.

TABLE OF CONTENTS

TABLE OF FIGURES

ABBREVIATIONS

ΔΔG	Difference between two ligand binding energies
B-B1	Bradykinin receptor B1
$complex_{myr}$	Myristate and I-FABP complex
$complex_{palm}$	Palmitate and I-FABP complex
ΔG	Free binding energy
FA	Fatty acid
FABP	Fatty acid binding protein
FEP	Free energy perturbation
H	Shannon entropy
I-FABP	Intestinal fatty acid binding protein
ILBP	Ileal lipid binding protein
L-FABP	Liver fatty acid binding protein
LIE	Linear interaction energy
MD	Molecular dynamics
TI	Thermodynamic integration

1 INTRODUCTION

The text book view of ligand-protein interactions is that rigid ligands lose less entropy whilst binding to a protein than flexible ligands do. Since the loss of entropy is related to the strength of the binding, then rigid ligands that lose less entropy are expected to bind to proteins in a stronger way than flexible ligands all else being equal (Murray and Verdonk, 2002, Chang et al., 2007, Searle and Williams, 1992, Mann, 2008). Thus, more flexible ligand should have a weaker affinity for the target protein. Nevertheless, flexible fatty acids (FA) showed a strong affinity for fatty acid binding proteins (FABPs). It has been reported that myristate and palmitate binding constants (K_i) to intestinal-FABP (I-FABP) are 0.041 and 0.024 μM, but the K_i of more rigid ligands, e.g. nitrazepam and bezafibrate are 2000 and 33 μM (Velkov et al., 2007). A smaller value of K_i indicates stronger ligand-protein interaction, thus, the binding strength to I-FABP can be ranked as: nitrazepam < bezafibrate < myristate < palmitate.

Based on the common dogma a suitable strategy for drug development is a restriction of conformational flexibility (Mann, 2008). Thus, the strong binding of flexible FA to FABPs is perplexing and suggests that both the common dogma and the binding of palmitate and myristate to FABPs are worthy of further consideration.

1.1 FLEXIBILITY AND ENTROPY

1.1.1 Flexibility

Molecule flexibility is defined by a molecule's ability to adopt a number of different shapes or conformations. High molecular movements over time describe very flexible molecules, while molecules restricted in intra-molecular movements are rigid. Molecular motions can be distinguished into translational, rotational and vibrational motions (Bender, 2003).

Translational movement of molecules occurs as a whole molecule movement which results in a change of position, in a space (Figure 1.1 (a)). Thus, translational molecular movements do not contribute to molecule flexibility. Nevertheless, restriction of translational motion does cause a loss of entropy (Murray and Verdonk, 2002).

Molecular rotation is the rotation of the whole molecule around the principal axis of the molecule (Figure 1.1 (b)). It does not have any contribution to the molecular flexibility whereas intra-molecular motions cause the biggest conformational change. Intra molecular rotations in an atom occur when four or more atoms are joined to each other linearly in the molecule, e.g. A-B-C-D, the B-C bond forms an axis of rotation, and there is a circular movement of atom D and A bond B-C (Figure 1.1 (c)). Thus, the bond B-C is referred to as a rotational bond.

Intra-molecular rotations in a molecule occur freely if the rotatable bond is single. When atom B and C are joined by a double bond movement of atom D is limited. If any external force induces movement of atom D, the double bond between atom B and C breaks (Housecroft and Constable, 2006). An increase in the number of rotatable bonds in a molecule increases the overall motions available to that molecule, i.e. - it becomes more flexible.

Vibrational motions consist of bond stretching, angle bending, rocking and libration (Figure 1.1 (d)). Vibrational movements occur as small movements of the atoms in the molecule, they thus have a small contribution to a molecule's conformational change and flexibility. Bond stretching is based on atomic vibration, which causes a change of the bond length during the interaction between two joined atoms. Vibrational motions of a bond occur at low amplitude since larger atomic motions cause the bond to break. Larger motions are observed for angle bending. When a molecule consists of three atoms, which are joined to each other, e.g. A-B-C, the angle bending occurs as the A and C atoms move toward and away from each other (Leach, 2001). Rocking, which is also called asymmetric stretching, is when at the same time atom A moves from atom B atom C moves towards atom B and vice versa (Housecroft and Constable, 2006). Librations occur as motions in atoms A and C, where they oscillate slightly backwards and forwards around the x, y or z axes.

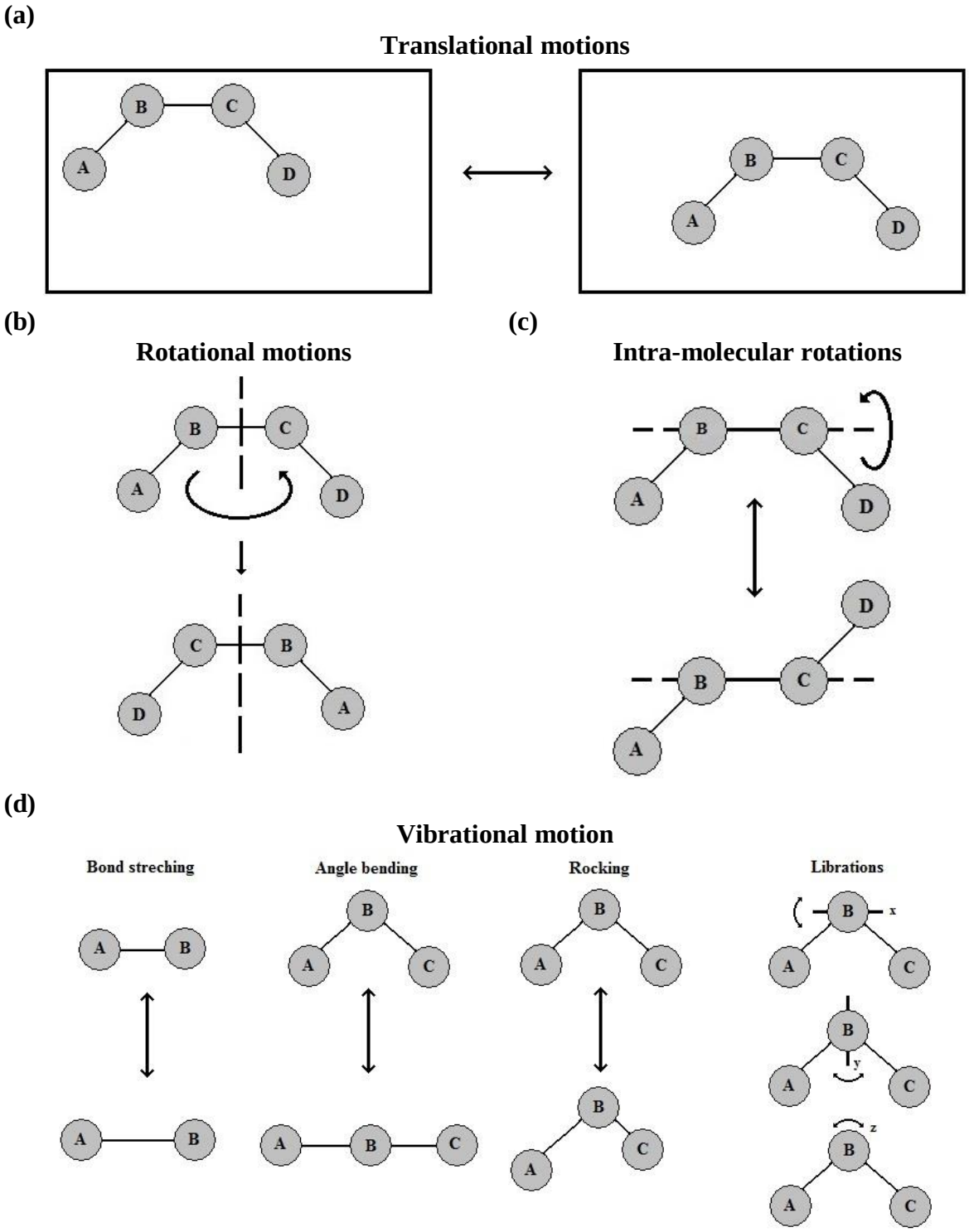


Figure 1.1 Molecular motions (a) Translational motion, where molecule A-B-C-D move from one place in a space to another one (b) Rotational motion around the principal axis of the molecule (c) intra-molecular rotations: rotation of atom D around the axis formed by the bond B-C (d) Vibrational motions: bond stretching, angle bending, rocking and librating. (reproduced from Leach, 1996, Chaplin, 2010)

1.1.2 Entropy and flexibility

Changes in dihedral angles, via rotation of atoms around single bonds, allow structural changes in a molecule. Moreover, high structural changes of molecules are characterized by their flexibility. Therefore, the number of rotatable bonds in a molecule contribute to the entropy change in different processes such as molecular association in solutions and fusion (Searle and Williams, 1992) as well as ligand-protein interactions (Searle and Williams, 1992, Murray and Verdonk, 2002, Chang et al., 2007).

During fusion of *n*-alkanes, *n*-alkyl carboxylic acids and *n*-alkyl methyl ketones an increase in the number of rotatable bonds increases the fusion entropy of these poly-organic compounds. For instance, data for odd *n*-alkanes fit with the equation $\Delta S = 15.5 + 1.6n$ (where n is the number of rotatable bonds) whereas even *n*-alkanes fit with the equation $\Delta S = 10.2 + 3.5n$ (Searle and Williams, 1992). Therefore, the gain in flexibility contributes to the gain in entropy during fusion of poly-organic compounds. Thus, flexible ligands have more entropy to lose while binding to a protein.

1.1.1.1 Ligand flexibility and entropy

Pharmaceutical drugs act as ligands. They bind to a protein and inhibit or activate cell signaling pathways with the aim to cure, treat or prevent a disease. A pharmaceutical drug with a low binding affinity to a target protein may be replaced by a competing ligand (Nelson and Cox, 2000). Thus, strong protein and ligand binding is relevant in drug design. Therefore, drug discovery studies are based on increasing the ligand binding affinity. Ligand binding affinity can be improved by making the enthalpy change more negative, the entropy change more positive or by a combination of both (Velazquez-Campoy 2001).

Our project is to investigate why flexible FAs bind strongly to I-FABP. A molecule's flexibility is based on its ability to adopt a number of conformations. Thus, we are interested

in the relation between the conformational entropy change and ligand flexibility. The entropy change is based on the changes in solvation entropy and conformational entropy. Thus, an increase in the entropy of binding can be provided by an increase in the ligand hydrophobicity as well by altering ligand geometry with the aim to minimize the loss of conformational entropy upon binding (Velazquez-Campoy 2001).

Chang et al (2007) studied the relation between the entropy change (-TΔS) and ligand flexibility in amprenavir binding to Human immunodeficiency virus (HIV) protease. Amprenavir is a flexible ligand and whilst binding to a HIV protease it loses configurational entropy ($-T\Delta S_{config}$ = 26.4 kcal/mol). This loss of entropy is based on the loss of vibrational entropy ($-T\Delta S_{vib}$ = 24.6 kcal/mol) while the loss of conformational entropy ($-T\Delta S_{conf}$) is only 1.8 kcal/mol. Vibrational entropy includes the entropy associated with rotation and translation in space and the motions associated with torsional degrees of freedom, which provides almost as much entropy change as the rotation and translational changes while change of conformational entropy depends on the number of the stable conformers.

Chang et al (2007) also calculated the loss of entropy while restricting amprenavir bond length and angles. By restricting bond length only the loss of the entropy stayed unchanged ($-T\Delta S_{config}$ = 26.3 kcal/mol) whilst by fixing all bond angles caused it to drop down to 24.7 kcal/mol. Conformational restriction of HIV protease strongly affect the loss of conformational entropy. The decrease in the loss of total entropy is due the decrease of the loss of conformational entropy.

The common dogma states that rigid ligands that lose less entropy are expected to bind to proteins in a stronger way than flexible ligands all else being equal. Thus, ligand binding affinity to a protein could be improved by an introduction of the rigidity in the ligand. The

introduced rigidity in HIV amprenavir more likely affects the loss of conformational entropy not the loss of vibrational entropy.

It might be that flexible FA loses a negligible amount of conformational and torsional entropy upon binding to a FABP. Thus, there is not a large change in the loss of conformational entropy between free and artificially restricted FA upon binding to FABP. This is what shall be investigated here.

1.1.1.2 Protein flexibility and entropy

Proteins are dynamic macromolecules. They may unfold, fold, bind and dissociate from a ligand. Entropy change accompanies all of these dynamic processes. Protein and ligand binding is usually associated with the loss of their motions. Nevertheless, the complex formed may allow for new vibrational modes within it (Boehr et al., 2006). For example the dynamics of the ligand-protein topoisomerase I complex are higher than it is for free protein (Yu et al., 1996, Boehr et al., 2006). Protein dynamics can stay unchanged as observed in the binding of R-lytic protease (Davis and Agard, 1998, Boehr et al., 2006).

Flexible proteins such as ion channels, nuclear hormones and transporters are able to bind a variety of ligands. Thus, protein flexibility may be essential for their functions. Moreover, reviewing data about protein and ligand complexes it was found that structurally different ligands may bind the same protein in different conformations, e.g. for drug-like molecule binding to therapeutically important receptors (Teague, 2003).

Flexibility in a protein not only allows the ability to bind a variety of ligands at different protein sites, but additionally a protein may be flexible at a binding site. For instance, HIV reverse transcriptase can be inhibited by structurally different ligands Efavirenz, Nevirapine, UC-781 and Cl-TIBO binding at the same site (Ren et.al, 2000). The structure of these ligands is presented in figure 1.2. Thus, it might be that flexibility of a

protein is crucial for the strong binding of a rigid ligand, while flexible ligands may better adapted for the binding to a rigid protein that do not have the right conformation.

The hydrogen-bonding network in ileal lipid binding protein (ILBP) is weaker than in FABPs indicating that ILBP is more flexible. This unusual flexibility of ILBP allows for the binding of the large and rigid bile acid (Lucke et al., 1996). Chemical shifts of protons in the ligand and protein complex caused by FA and the more rigid chenodeoxycholic acid binding with ILBP were observed using nuclear magnetic resonance (NMR) spectrometry

A 14% difference in proton chemical shifts was observed when chenodeoxycholic acid binds to the protein, but for FA upon its binding to the protein it was just 1%. Bigger chemical shifts observed for the chenodeoxycholic acid binding to ILBP may be due to a conformational change or due to the presence of ligand atoms next to the protein (Lucke et al., 1996). Although, the greater interaction between ILBP and chenodeoxycholic acid was reported, there was no information on whether it causes stronger binding. However, it might be that due to flexibility proteins are likely to bind rigid ligands more strongly than flexible ones.

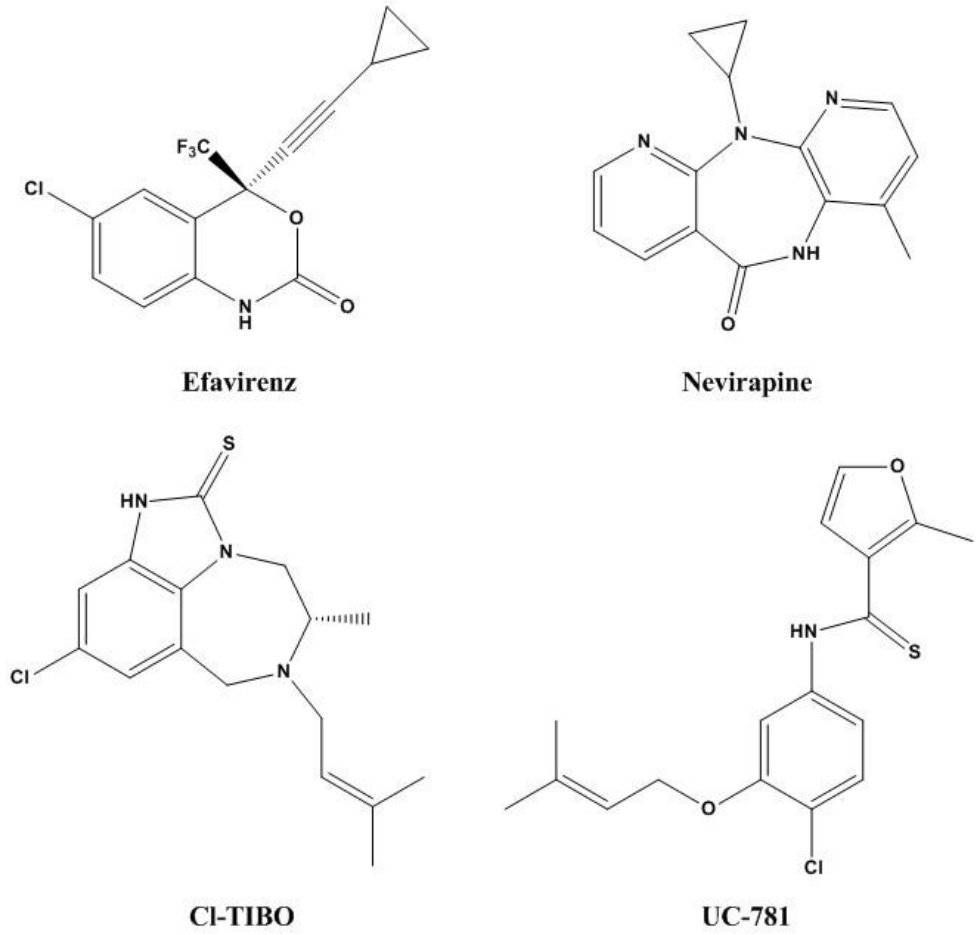


Figure 1.2 Structures of ligands able bind to HIV reverse transcriptase at the same binding site (reproduced from Ren et.al, 2000).

1.2 THE ROLE OF FLEXIBILITY IN THE THERMODYNAMICS OF LIGAND AND PROTEIN INTERACTIONS

1.2.1 Evidence supporting ligand flexibility being detrimental to ligand-protein interactions

Common dogma states that rigid ligands bind more strongly to a protein than a flexible ligand does. The flexibility of a ligand can be determined by the number of rotatable bonds per molecule. Therefore, common dogma emphasizes the importance of rotatable bonds in the control of ligand binding affinity, and this is reflected in Medicinal chemistry text books e.g., (Mann, 2008). Mann (2008) reports about the introduction of rigidity in such ligands as phenylalanine, histamine and milnacipran analogs, and discussed many examples, where an increase of ligand rigidity by the removal of rotatable bonds, contributes to an increase in binding affinity of ligands. An example is given in figure 1.3. The restriction of compound **1** at the benzyl position increases its rigidity, thereby the newly formed compound **2** has a higher affinity for the Bradykinin receptor B1 (B-B1). The K_i of compounds **1** and **2** are 520 nM and 380 nM, respectively. Another example is the increasing of rigidity in compound **3** to become compound **4** and subsequently compound **5**, where the K_i for binding to B-B1 is 382 nM, 0.24 nM and 0.77 nM, respectively (Figure 1.3). Therefore, compounds **4** and **5** have a greater affinity for B-B1 than does compound **3** (Mann, 2008).

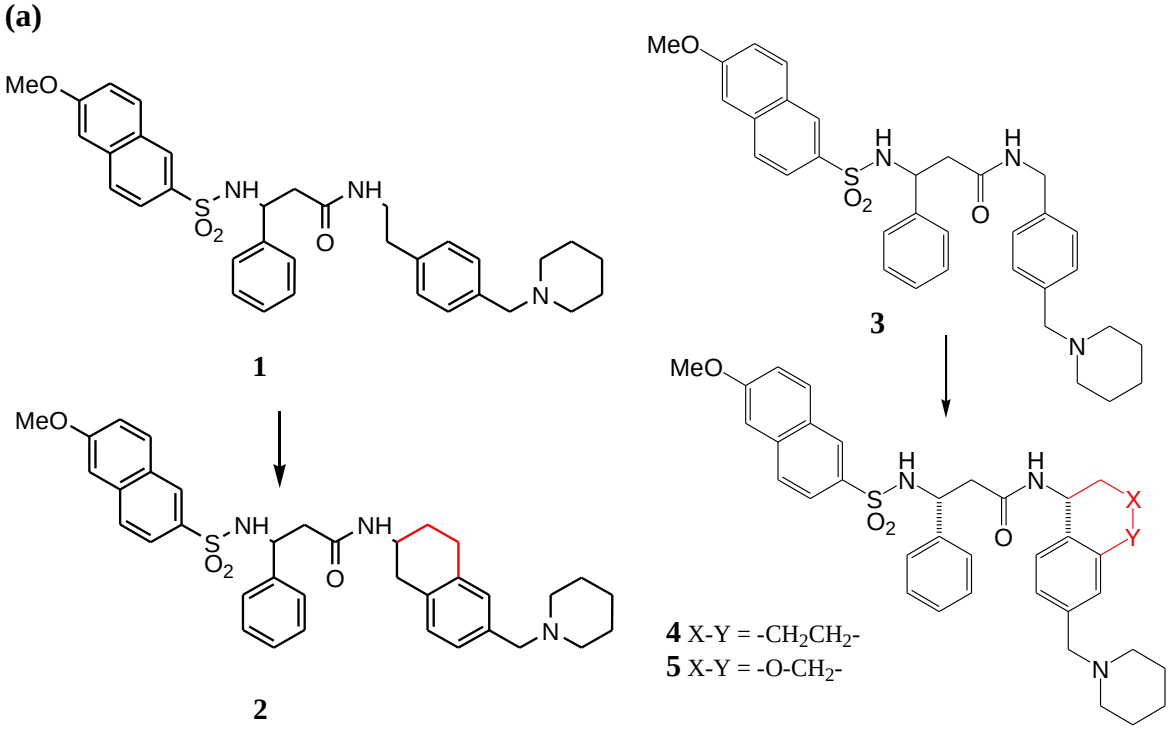
The removal of rotatable bonds often occurs via the introduction of one or more extra heavy atoms (non-hydrogen). The increase in heavy atoms of a ligand up to 15 increases the free binding energy by on average 1.5 kcal/mol (6.3 kJ/mol) per heavy atom (Kuntz et al., 1999), increasing the binding affinity of the ligand. The addition of an extra heavy atom in the large ligands (> 15 heavy atoms) gives a small increase in free binding energy. Therefore, it is not clear if the improved affinity of a compound with restricted torsions is due to extra heavy

atoms added to restrict the torsions, or due to the decreased flexibility. Ligand efficiency is the free binding energy per heavy atom (Abad-Zapatero, 2007), thus comparing the ligand efficiency of different ligands provides a way to control for the variation in the number of heavy atoms.

Binding free energy is calculated using the formula: $\Delta G=RTlnK_i$, where R is the universal gas constant, T is the temperature and K_i is the inhibitory constant, a measure of binding affinity. Ligand efficiency is obtained by dividing the calculated ΔG value by the number of heavy atoms. Binding free energy increases with each heavy atom introduced in a molecule. Thus, similar ligand efficiencies would mean that ligands bind stronger due to heavy atom introduction. While a higher value of ligand efficiency would indicate that an increase in ligand binding affinity is not likely to be just due to the introduction of a heavy atom, but to some other effect such as the molecule becoming more rigid as well.

The number of rotatable bonds is decreased by 2 when compound **1** is restricted to form compound **2**. Nevertheless, compounds **1** and **2** have similar ligand efficiencies 1.24 kJ/mol and 1.20 kJ/mol, respectively (Figure 1.3). Although, the decrease in the binding constant supports the dogma when compound **1** becomes restricted to form compound **2**, compound **2** may bind more strongly to B-B1 due to an increase in heavy atoms in the molecule, not due to the increase in rigidity. Ligand efficiencies for compound **4** and **5** are approximately 1.2 times greater than it is for compound **3** (Figure 1.3). Therefore, stronger binding of **4** and **5** to B-B1 is likely to be due to the reduction in the number of rotatable bonds.

(a)



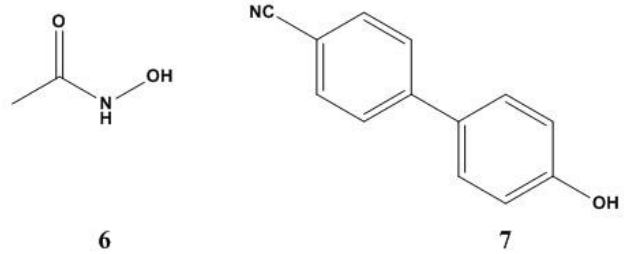
(b)

Ligand	K_i, nM	ΔG, kJ/mol	The count of heavy atoms	Ligand efficiency (ΔG/the count of heavy atoms), kJ/mol	The count of rotatable bonds	The count of Rotatable bonds/the count of heavy atoms
Compound **1**	520	-52.05	42	1.24	12	0.29
Compound **2**	380	-52.81	44	1.20	10	0.23
Compound **3**	382	-52.80	41	1.29	11	0.27
Compound **4**	0.24	-70.75	44	1.61	10	0.23
Compound **5**	0.77	-67.91	44	1.54	10	0.23

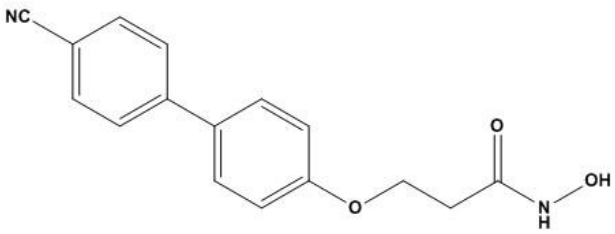
Figure 1.3 Drug design making ligands more rigid by the introduction of extra bonds (a) Design of new Bradykinin ligands. (b) Calculated Bradykinin ligand efficiencies (Ligand structures are reproduced and K_i values are taken from Mann, 2008).

A fragment-based drug discovery approach was introduced in the mid-1990s. However the technique first time was described earlier by Page and Jenks (1971). This technique is based on the joining two fragments which have an affinity to bind a protein in different but adjacent binding sites. Two fragment joining result in entropy gain because when joined fragment binds to a protein it loses less entropy than individual fragments do. Thus, it was reported that the affinity of the joined fragment will usually be greater than the sum of two fragment affinities (Murray and Verdonk, 2002). An example is given in figure 1.4; compounds **6** and **7** are joined to form compound **8**, while compounds **9** and **10** are joined to form compound **11**. The K_i of joined compounds **8** and **11** is much greater than it is for the individual fragments.

(a)

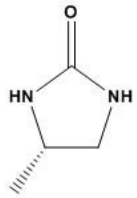


6 7

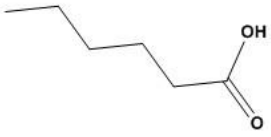


8

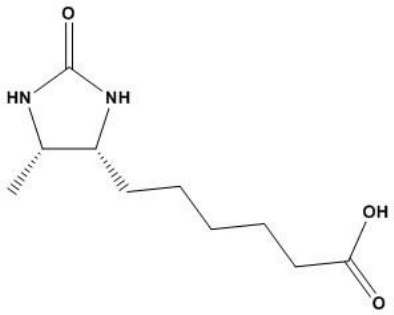
(b)



9



10



11

(c)

Ligand	Molecular mass	K_i, μM	ΔG, kJ/mol	The count of Heavy atoms	Ligand efficiency (G/the count of heavy atoms), kJ/mol	The count of rotatable bonds	The count of Rotatable bonds/the count of heavy atoms
Compound **6**	75	17000	-9.92	5	1.98	2	0.40
Compound **7**	195	20	-26.34	15	1.76	2	0.13
Compound **8**	282	0.025	-42.62	21	22.03	6	0.29
Compound **9**	100	34	-25.05	7	3.58	0	0.00
Compound **10**	116	260	-20.10	8	2.51	4	0.50
Compound **11**	214	0.00000041	-69.45	15	4.63	6	0.40

Figure 1.4 Drug design making ligands more rigid by joining two fragments together (a) Joining compound 6 and 7 is formed new ligand – compound 8. Compound 6, 7 and 8 binds to stromelysin. (b) Joining compound 9 and 10 is formed new ligand – compound 11. Compound 9, 10 and 11 binds to avidin. (c) Calculated ligand efficiencies. (Ligand structures are reproduced and K_i are taken from Murray and Verdonk, 2002)

1.2.2 Ligand flexibility may be favorable

Ligand conformational restriction is a basic principle in drug-design. Nevertheless, it was reported that ligand flexibility may be favorable as flexible ligand can adapt and bind to mutated proteins.

Protein mutation can cause binding site conformational changes making a protein become drug resistant (Cai and Schiffer, 2010). Thus, a mutated protein may not bind a rigid ligand that has a high affinity for the wild type protein (Velazquez-Campoy et al., 2001). Moreover, multiple mutations often affect the binding of ligands even when the mutations are distant from the binding site (Lassila, 2010, Genoni et al., 2010). Therefore, flexible ligands are thought to be more favorable as they are able to adapt their conformation so as to bind the mutated protein. As an example, KNI-272 has a high affinity for HIV-1 protease. The K_i is 16 pM. Since KNI-272 is rigid, it is not able to adapt to the protein mutation and the binding affinity to the HIV-1 protease mutant V82F/I84V is 550 times smaller. Flexible KNI-764 ligand has a slightly lower affinity than KNI-272 for HIV-1 protease; the K_i is 36 pM. Due to its flexibility it can adapt to the conformational change in the binding site of mutant V82F/I84V. The affinity of ligand KNI-764 for V82F/I84V is decreased just 26 times (Velazquez-Campoy et al., 2001). KNI-272 and KNI-764 structures are represented in figure 1.5.

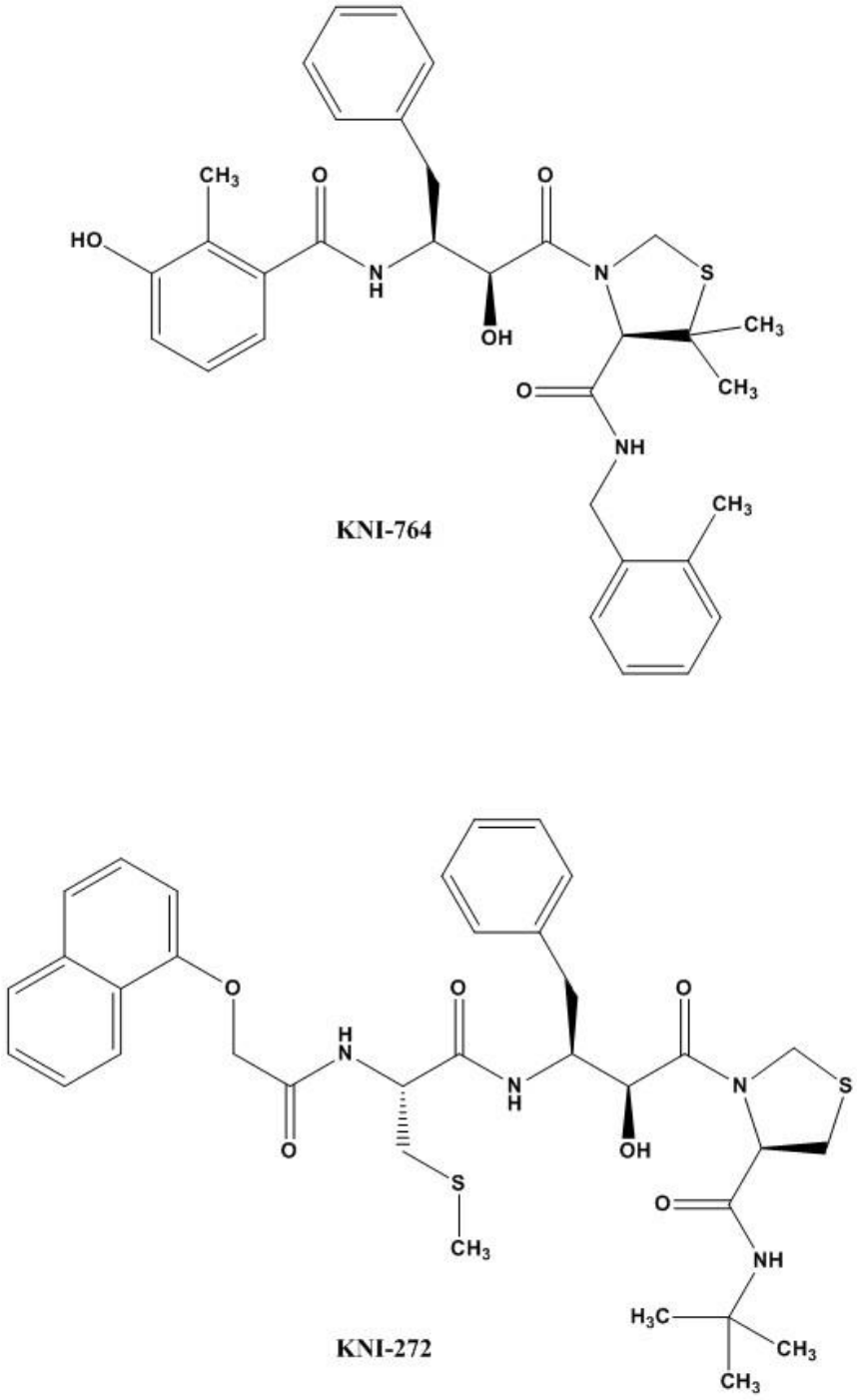


Figure 1.5 Structures of ligands able bind to HIV-1 protease (reproduced from Velazquez-Campoy et al., 2001).

1.2.3 Evidence for exceptions, where flexible ligands bind stronger to a protein than rigid and larger ligands do

Common dogma states that larger and more rigid ligands bind more strongly to a protein. The evidence supporting the dogma described above showed that it is enough to make a ligand larger or more rigid to increase its binding affinity. Thus, drug-design is based on the introduction of extra rigidity. Nevertheless, in review Teague (2003) reported that there are a large number of flexible ligands used as drugs. For example, flexible zopolrestat and tolrestat bind to aldose reductase, while flexible tamoxifen and raloxifen bind to oestrogen receptor.

It was experimentally determined that flexible FAs bind stronger to I-FABP than larger and more rigid ligands do (Velkov et al., 2005, Velkov et al., 2007). The K_i for palmitate and myristate are $\geq$ 1000 times higher than K_i of rigid nitrazepam and bezafibrate. Calculated ligand efficiencies for palmitate, myristate, nitrazepam and bezafibrate bound to I-FABP show that flexible FAs have higher ligand efficiency as well as the count of rotatable bonds per heavy atom. Thus, it might be that flexibility of FA is relevant for strong binding to I-FABP and it is an exception to the common dogma.

Palmitate differs from myristate by two extra heavy atoms. Thus, it has extra rotatable bonds and is more flexible than myristate. K_i values indicate that flexible palmitate binds more strongly to I-FABP than myristate. Nevertheless, the calculated ligand efficiencies show that the two additional heavy atoms decrease the free binding energy per heavy atom. Therefore, the two additional heavy atoms are likely to be the reason for the palmitate stronger binding to I-FABP, not increase in flexibility.

Another example is the strong binding of oleate to liver-FABP (L-FABP) (Chuang et al., 2008). The difference between K_i values for oleate and the more rigid and larger ligands is not as significant as it was in the case of myristate and palmitate binding to I-FABP. For

example, the K_is of oleate, ibuprofen and fenofibric acid are 0.18 μM, 47.6 μM and 0.334 μM, respectively.

Intriguingly there are larger and more rigid ligands that bind to L-FABP stronger than oleate. They are progesterone and fenofibrate, whose structures are represented in figure 1.6. The calculated ligand efficiencies for oleate, progesterone and fenofibrate are very similar while the count of rotatable bonds in oleate is much bigger (Figure 1.7).

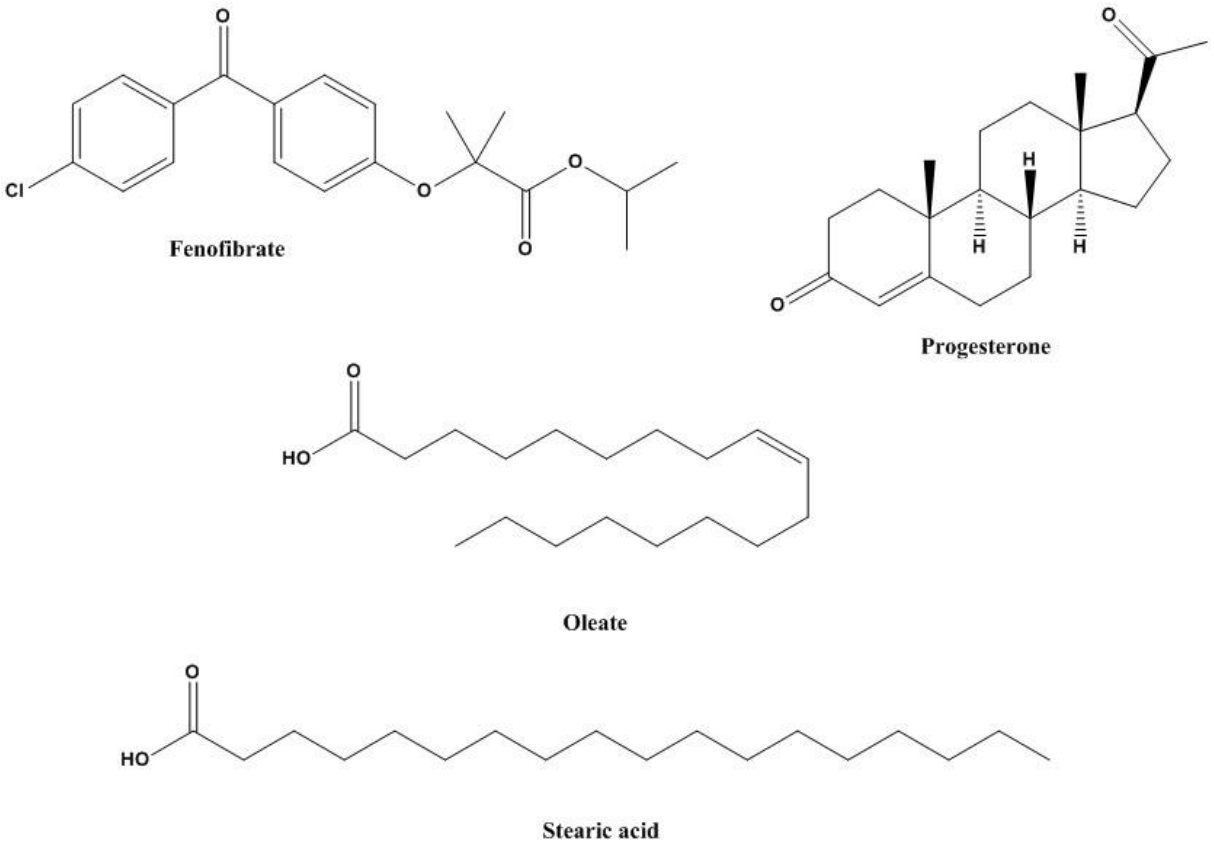


Figure 1.6 Structures of ligands (reproduced from Chuang et al., 2008).

(a)

Ligand	K_i, µM	ΔG, kJ/mol	The count of heavy atoms	Ligand efficiency (G/the count of heavy atoms), kJ/mol	The count of rotatable bonds	The count of Rotatable bonds/the count of heavy atoms
Nitrazepam	2000	-15.13	21	0.72	2	0.10
Bezafibrate	33	-25.13	25	1.01	8	0.32
Myristate	0.041	-41.42	16	2.59	12	0.75
Palmitate	0.024	-42.72	18	2.37	14	0.78

(b)

Ligand	K_i, µM	ΔG, kJ/mol	The count of heavy atoms	Ligand efficiency (G/the count of heavy atoms), kJ/mol	The count of rotatable bonds	The count of Rotatable bonds/the count of heavy atoms
Oleate	0.18	-37.82	20	1.89	15	0.75
Progesterone	0.027	-42.43	23	1.84	1	0.04
Fenofibrate	0.024	-42.72	25	1.71	7	0.28
Fenofibric acid	0.334	-24.23	22	1.15	5	0.23
Ibuprofen	47.6	-36.31	21	1.65	3	0.14

Figure 1.7 Ligand efficiency (a) Calculated ligand efficiency, when they bind to I-FABP. (b) Calculated ligand efficiency, when they bind to L-FABP (K_i are taken from Velkov et al, 2007, Chuang et al, 2008).

1.3 I-FABP: A MODEL SYSTEM FOR EXPLORING THE ROLE OF FLEXIBILITY IN LIGAND BINDING

Common dogma states that rigid ligands have a higher affinity when binding to a protein, due to a smaller loss of entropy. Calculated efficiencies for Bradykinin ligands show that increase in ligand binding affinity may be due to the increase in the number of heavy atoms. Thus, it is questionable if the ligands' strong binding is due to their rigidity. Moreover, some flexible ligands appear to bind to protein more strongly than rigid ligands, something seen in the I-FABP system. Therefore, it is relevant to investigate the role of ligand flexibility during the binding of flexible FAs to I-FABP.

FAs are small size (150-450 Da) structural components of the cell's phospholipid membrane and undertake an important part in the cell's metabolic processes. Nevertheless, the presence of free FAs in the cytoplasm is unwanted due to their detergent-like properties (Tsfadia et al., 2007, Levin et al., 2010, Levin et al., 2009, Friedman et al., 2006, Woolf and Tychko, 1998). Therefore, efficient transfer of free FAs in the cytoplasm to the cell's structural elements is needed. This is performed by FABPs, which have a high affinity to bind FAs. FABPs are classified by their tissue of origin, e.g. I-FABP, L-FABP and muscle-FABP (M-FABP) (Friedman et al., 2006, Levin et al., 2009, Tsfadia et al., 2007, Velkov et al., 2005).

The most important finding is to be stressed here is that FAs such as myristate and palmitate bind more strongly to I-FABPs than more rigid ligands do (Velkov et al., 2005). This observation is opposite to the dogma. Palmitate is more flexible than myristate, due to its higher number of rotatable bonds. However, the calculated ligand efficiencies show that the stronger binding of palmitate is due to the increase in the number of heavy atoms. Thus, it seems that flexibility of FA is neither advantage nor disadvantage for their strong binding to

FABP. Therefore, it is questionable what kind of role plays FA flexibility in their strong binding to I-FABP.

The effect of ligand flexibility whilst binding to proteins can be determined by an investigation of myristate binding to I-FABP by computational modeling of the binding. As flexibility is related to the conformational change of a molecule, it is important to observe conformational contribution of myristate while binding to I-FABP. It could be true that common dogma does not have any exemptions and the strong binding of flexible FAs to FABPs is putative. On the other, hand it could be that ligand flexibility is somehow critical for binding.

1.4 COMPUTATIONAL APPROACH

Computationally it is possible to change parameters that define the flexibility of a molecule without changing the number of heavy atoms or its geometry. A computational approach requires the choice of an appropriate force field, one that demonstrates the appropriate behavior for the bind of myristate to I-FABP (i.e. similar to the known experimental behavior of the system). To determine if the force field is correctly chosen, a calculation of the difference between the ΔG of myristate and palmitate (ΔΔG) on binding to I-FABP can be made and compared to experiment. Further, the conformational preference of myristate in solution and bound to I-FABP can be calculated by molecular dynamic (MD) simulations. This allows the determination of which myristate rotatable bonds become restricted when it binds to I-FABP. Furthermore, by calculating ΔΔG both when myristate is flexible and artificially rigid, the effect of flexibility on the binding energy can be quantified.

1.4.1 Force fields and molecular dynamics

Molecular dynamic simulation is the computational simulation of atom and molecule movements (Alder and Wainwright, 1959). Various protein motions occur over a broad timescale (Figure 1.8) (Boehr et al., 2006).. The time to simulate these motions can last from hours to several years. For example, whilst simulating ligand and protein binding, it is important to take into account the timescale of the molecule's vibrational motions (in picoseconds (ps) and nanoseconds (ns) time) and larger configurational changes (in microseconds (μs) and milliseconds (ms) time). Protein simulations of timescales of μs unusual but becoming more common and ms simulations are not currently practical for anything but short peptides. According to the computational time used for our project to simulate 1 μs of ligand motions on an eight core processor in real time is the equivalent of 0.5 years, whilst ligand and protein interactions will take around a year of computational time.

Nevertheless, MD simulations can be performed for a short computational time and subsequently extended in length to check for convergence of the simulated motions or thermodynamic properties, e.g. an initial simulation of 10 ns, with the behavior observed in this simulation confirmed by extension of the simulation to 20 ns, 50 ns etc...

To perform MD, a topology, a co-ordinate and a simulation input file are needed. The topology file contains information about the atom connectivity in a molecule, their potential energy and the interaction forces. The potential energy and interaction forces are defined by the force field. The co-ordinate file defines the initial atom positions. It will include the position of the solvent atoms if they are present in the system. The simulation input file defines simulation parameters such as temperature and pressure. Following minimization of the energy of the initial coordinates, to remove high energy strains from the system and an initial equilibration only then is the simulation system ready for productive MD. Energy minimization allows for the initial structure to adjust to the constrains of the force field, but equilibration equilibrates kinetic and potential energies through the simulation system.

In a molecular modeling context a force field is a mathematical function, which evaluates the potential energy of the atoms in the system. Force fields for computational simulations are derived from quantum mechanical calculations and from experimental data. Based on the required representation and an available amount of computational time, force fields can be divided into three categories. Firstly, “all atom” force fields, where energetic interactions are represented for all atoms. Secondly, “united atom” force fields, where all atoms in a terminal methyl group are represented in the force field as if they were one atom. Thirdly, “coarse grained” force fields that provide a much cruder representations, e.g. an amino acid represented by a spherical particle for the backbone and a spherical particle for the

side chain (Ponder and Case, 2003). Moreover, by simplifying the contribution of some atoms this makes the computation of system's energy less time-consuming.

The first force fields for the study of macromolecule dynamics were developed and designated as AMBER (Cornell et al., 1995), CHARMM (Brooks et al., 1983) and GROMOS (Scott et al., 1999). AMBER force field was originally developed for protein and DNA study. CHARMM, besides being used in macromolecule research, was also developed in small molecule research. However, GROMOS is used for the study of biomolecular systems and has two versions – A and B. The A version is intended for the simulation of aqueous or apolar solutions of proteins, nucleotides and sugars. The B version is intended for the simulation of isolated molecules in their gas phase. Besides these three, there are many more varieties of force field with varied applications, i.e. more or less specific (Ponder and Case, 2003).

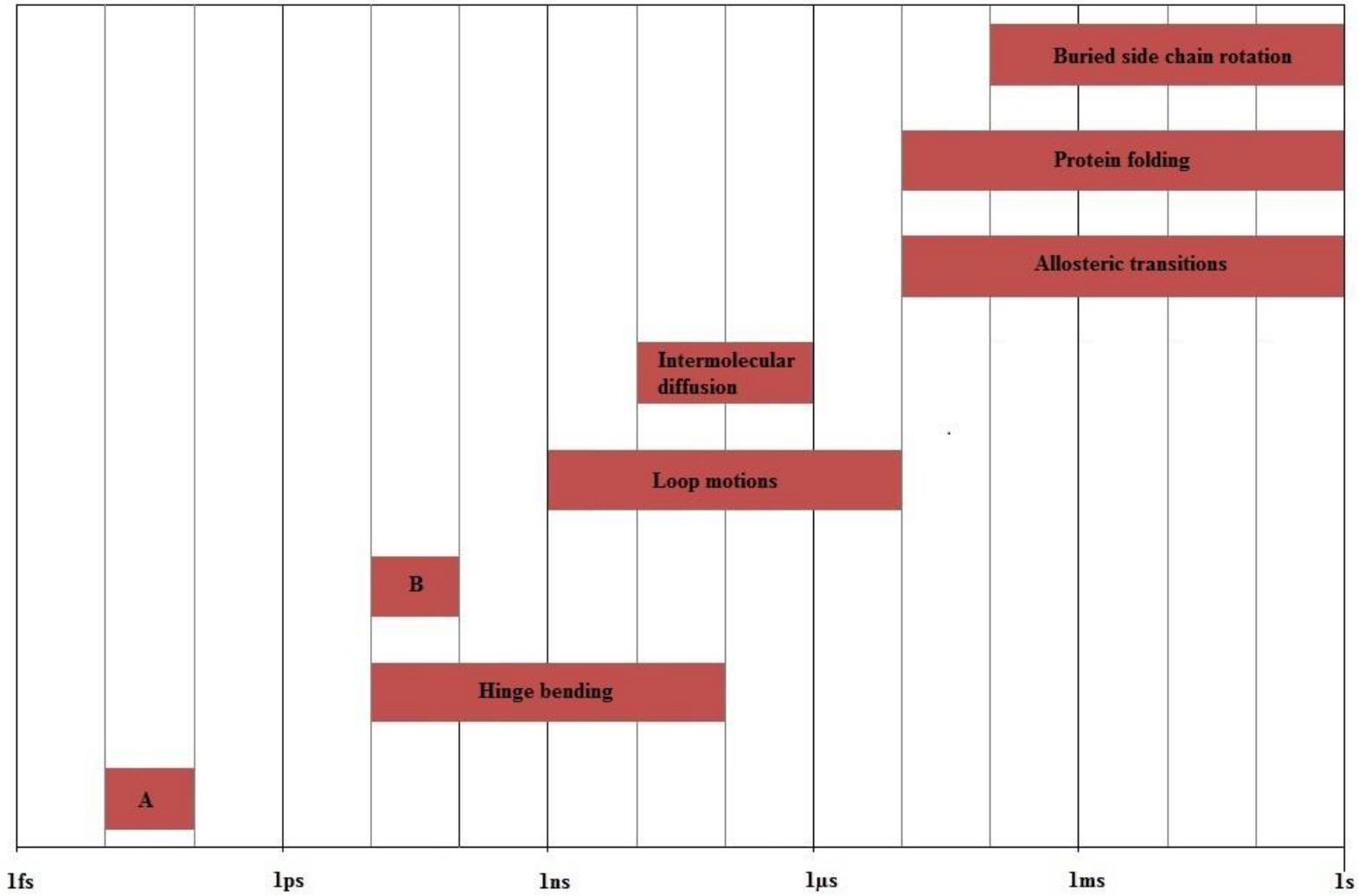


Figure 1.8 Timescale of typical protein motions A – bond vibration and B – surface side chain rotations (reproduced from Boehr et al., 2006).

1.4.2 Calculation of binding free energy

The calculation of free energy is based on the thermodynamic cycle of the binding of proteins and ligands. Computationally we calculate the conversion for the solvated free ligand X into the solvated free ligand Y and for the conversion of ligand X into ligand Y while they are bound to the solvated protein P. Water is usually used as a solvent for these simulations. Computational mutation of the ligand gives two changes of free energy, which are calculated for the unbound state (ΔG_U) and for the bound state (ΔG_B). The difference in free energy of the ligands ($\Delta\Delta G_{bind}$) is calculated by the formula shown in figure 1.9 (Michel et al., 2010).

For free energy computational calculations, several methods have been presented: free energy perturbation (FEP), thermodynamic integration (TI), linear interaction energy (LIE), molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) and molecular mechanics generalized Born surface area (MM-GBSA) (reviewed in Michel et al., 2010, Hayes et al., 2012). Considering about which method to use there is a need to make a choice between expenses and accuracy. Such methods as LIE, MM-PBSA and MM-GBSA were reported as trade-off and are successfully used. FEP and TI calculation methods are most accurate of these methods, although they are time consuming. Of these two TI has the advantage that it is easier to parallelize since, as described below, each lambda point used in the calculation is completely independent of every other lambda point.

The computational approach of FEP allows the calculation of the potential energy difference U_Y - U_X between two complexes during the conversion of the ligand X into the ligand Y. The conversion in the ligands is maintained throughout one or more intermediate states defined by a coupling parameter λ. The coupling parameter λ is used for the acquisition of free energy from intermediate states. Therefore, $\lambda = 0$ represents the ligand X, but $\lambda = 1$ represents the ligand Y. When λ is any value between 0 and 1, it represents an intermediate

state of ligand X changing to ligand Y (Michel et al., 2010). A simplified model is present in figure 1.9 (b). Stars X and Y represent two structurally different ligands. Those differences are illustrated by color changes in the figure. Therefore, stars A to F represent intermediate states in the ligand X mutation to Y. Moreover, those intermediate states do not represent any different molecule. Depending on λ value, they illustrate whether they are structurally more similar to ligand X or Y, but their structure does not have any real physical meaning in the calculation of binding free energy. The ΔG calculation is more accurate when a higher number of chosen intermediate states are used. Each state gives free energy ΔG, which is an averaged value of U_{λ}-$U_{\lambda-1}$ (where λ – 1 denotes the lambda value before the one under consideration) over numerous snapshots for given lambda value). The total value of all ΔG gives ΔG_B.

TI has a similar principle for calculating free energy ΔG_B, using the intermediate states. However, this method integrates the derivate of the potential energy for with respect to lambda point in the conversion of the ligand X into Y using equation:

$$\Delta G_B = \int_0^1 \langle \frac{\partial U(\lambda)}{\partial \lambda} \rangle_\lambda \, d\lambda$$

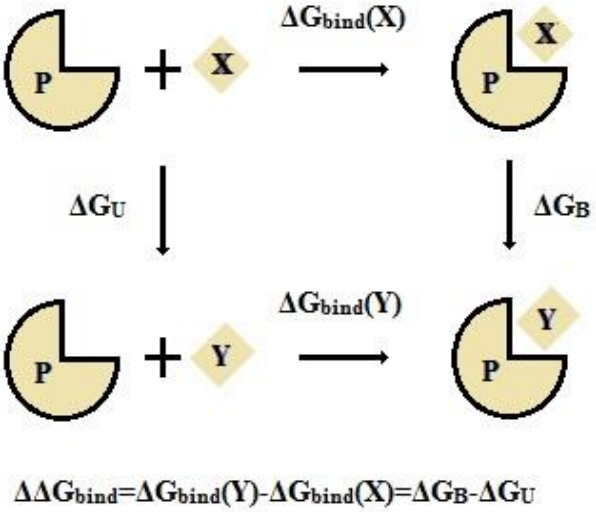
In the equation ΔG_B is the total free energy during ligand X conversion into Y for the ligand-protein complex, where $\langle \frac{\partial U(\lambda)}{\partial \lambda} \rangle_\lambda$ is the ensemble average in the ensemble corresponding to U(λ) (formula is adapted from Steinbrecher et al, 2007). TI has the advantage over FEP that it is easy to add additional lambda points, since the derivative depends only on the lambda value at hand, not on the preceding lambda value, which is the case in FEP.

FEP and TI are limited in usage due to ligand X cannot be converted into any different ligand Y. The ligands need to have similar structures to each other. The most accurate results

are obtained when X and Y ligands have small differences, of particular concern is that the ligands' difference in shape and size can cause different preferred locations at the binding site, which would represent very large perturbations to the system. FEP and TI give highly accurate results; however obtained results may differ from the experimental data (Michel et al., 2010, Steinbrecher et al., 2007, Cai and Schiffer, Michel et al., 2010). This can be overcome by altering the force field.

When investigating the gain or loss of atoms TIs are performed as a three-step simulation. In the first, the charge removal from the atoms to be lost is simulated. In the second the chosen atoms are transformed, so the ligand X became ligand Y. In the last step the new charge is added to the newly formed atoms. This three step procedure avoids the possibility that as the van der Waals forces are decreased with lambda, the charges on the atoms may become unrealistically close leading to excessively high interaction energies and thus problems with the simulation.

(a)



(b)

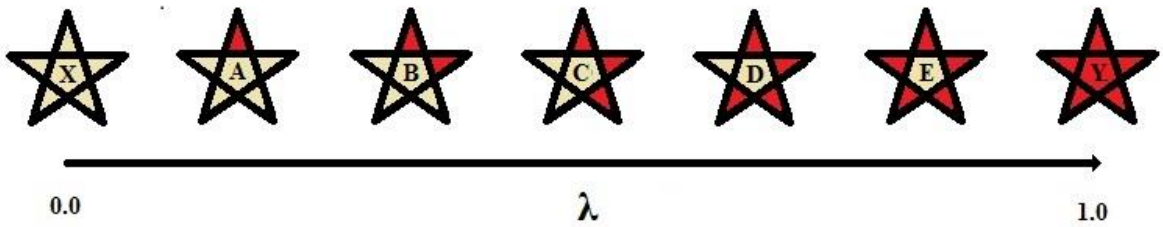


Figure 1.9 TI and free binding energy calculation (a) TI cycle to compute free energies of binding. P is a protein, X and Y are two different ligands, ΔG_U and ΔG_B are free binding energies for mutation of ligand at unbound and bound state, respectively. $\Delta G_{bind}(X)$ and $\Delta G_{bind}(Y)$ are free binding energies for ligand X and Y binding to a protein. $\Delta\Delta G_{bind}$ is the difference between ligand Y and X free binding energy. (b) Representative figure of λ value meaning. X and Y are ligands, A, B, C, D, E are intermediate states between ligand X and Y. Color difference between two stars identify how different or similar are two intermediate states between each other or from ligand X and Y. (Figures are reproduced from Michel et al, 2010).

1.4.3 Clustering

The conformational states of myristate can be observed by molecular dynamics (MD) simulations, which "produce trajectories of atomic positions (and optionally velocities and energy) as a function of time" (Shao et al., 2007). Data-mining techniques such as clustering sorts and groups the data enclosed in MD trajectories into similar clusters allowing investigate the preferable molecule's conformations.

There exist a couple of different clustering methods. Shao et al (2007) tested and compared clustering algorithms for single-stranded DNA molecules or DNA hairpin interactions with the minor groove binding drug DB226. The clustering algorithms tested include hierarchical, single-linkage edge joining, centroid-linkage, average-linkage, complete-linkage, centripetal and centripetal-complete, means, Bayesian, self-organizing maps and tree (COBWEB). The authors refer to each algorithm as having its own unique clustering abilities and none of the algorithms presented all possible abilities. Moreover, each algorithm is based on a comparison of different sets of results for the same molecule, but at different simulation times. Each algorithm has a different way of comparing the molecule. Therefore, each one gives varied result clusters. Shao et al (2007) found that grouping DNA molecules into 5 clusters to be the best for the observation of conformational change of DNA molecules, due to tendency to form single molecule clusters when molecules are grouped into > 5 clusters.

The average-linkage and hierarchical algorithms have an advantage as they can be used when the number of clusters required is not known. However, the hierarchical algorithm is highly sensitive to outliers and the average-linkage algorithm tends to produce small clusters. This tendency for the average-linkage algorithm can be minimized by performing clustering with different numbers of clusters. Myristate clustering into distinct sets will be performed using the average-linkage algorithm. Clustering by the average link algorithm can

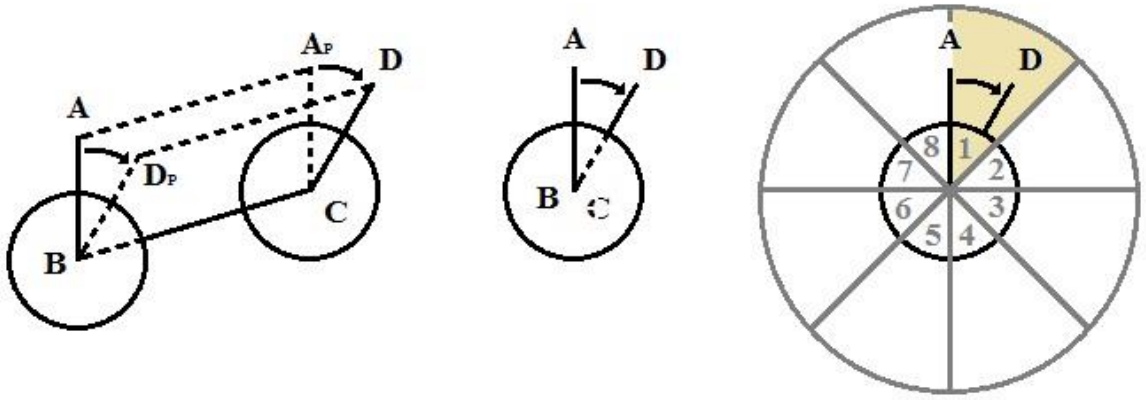
give some information about the increase in rigidity of myristate while it binds to I-FABP.

The dihedral angle is the angle or torsion between two planes. Thus, the dihedral angle for a sequence A-B-C-D is formed by planes $ABCA_P$ and D_PBCD, where A_P-C is the projection of bond A-B and D_P-B is the projection of bond D-C (Figure 1.10 (a.1)). The dihedral clustering principle is based on dividing a dihedral angle into bins (Case et al., 2010). For instance, if grouped in bins of eight, each bin differentiates the position of the atom D per 40º. Figure 1.10 (a.3) shows that atom D appears in the first bin. Therefore, the dihedral angle is between 0º and 40º. 180º is the maximum possible angle between bonds A-B and C-D (Figure 1.10 (b)). However, a dihedral angle between 320º and 360º would place it in the highest bin, bin 8. This bin is equivalent to 0º to -40º which we use later (Figure 1.10 (b)).

In the case of free myristate, it is expected to observe different occupancies of the bins representing different angles as myristate is not restricted and can, therefore, freely change its conformation. Occupancy of all the bins would represent the potential ability of free myristate to access all possible dihedral angle values for every aliphatic carbon torsion. Myristate bound to I-FABP might be expected to be more rigid than myristate free in solution. Therefore, fewer bins are likely to be occupied i.e. there is a more restricted number of dihedral angle values that are observed.

(a)

(a.1) **(a.2)** **(a.3)**



(b)

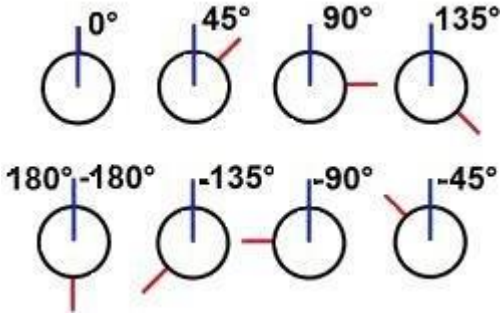


Figure 1.10 Dihedral angle (a) Dihedral angle and clustering bins, where (a.1) is the 3D structure of a chain A-B-C-D, atoms B and C are drawn as circles. The dihedral angle for a sequence A-B-C-D is formed by planes $ABCA_P$ and D_PBCD, where A_P-C is A-B bond projection and D_P-B is D-C bond projection. Dihedral angle is shown as an arrow. (a.2) Newman projection for the chain A-B-C-D. Atom C is behind atom B. Dihedral angle is shown as an arrow. (a.3) Clustering by dihedral angle into 8 bins (numbered in grey) (b) Newman projection of dihedral angles and their values.

1.4.4 The investigation system

Our computational model of myristate binding to I-FABP will be studied using the AMBER force field. TI simulations for the myristate change to palmitate will allow the calculation of the difference in the protein binding free energy of the two ligands. The data recorded during the MD simulation will be used for the comparison of the conformations of myristate that exist in its free and protein bound states. The comparison will be performed by the clustering of myristate conformational states using average-linkage and also by dihedral angle.

Combining MD simulations with clustering of myristate conformations, it can be determined whether myristate possesses any favored conformational state in solution or when it is bound to I-FABP. With the information about the conformational state of the free myristate and the bound myristate to I-FABP, it can be ascertained which bond torsions show changes of behavior upon binding with the protein. Further, using TI simulations, the free binding energy of myristate can be calculated, when myristate is both flexible and artificially rigid. This will give data to support whether flexibility of myristate plays an important role in its stronger binding to I-FABP than it was seen for rigid ligands.

1.5 THE AIM OF THE RESEARCH

Common dogma predicts that rigid ligands lose less entropy during binding, thus binding more strongly to protein than otherwise equivalent flexible ligands do. Nevertheless, experimentally it was found that flexible myristate and palmitate bind more strongly to I-FABP that rigid ligands do. Therefore in our research we examined I-FABP and myristate interaction using before described model with the aim of understanding the basis of stronger binding between flexible ligands and proteins.

An investigation of FABP and FA interactions is necessary as it is not clear is the flexibility of FA plays a role in the strong binding of FA and FABP. Moreover, this mechanism might work in other biological systems which include a high degree of conformational freedom. Thus, the aim of the research is to test if flexibility of myristate plays a role in myristate strong binding to I-FABP.

From a comparison of the calculated ligand efficiencies, it can be assumed that flexible palmitate binds stronger to I-FABP than less flexible myristate due to an increase in the number of heavy atoms. Therefore, it could be true that common dogma does not have any exemptions and flexible FA binding to FABP is just putative. Nevertheless, we are expecting that the flexibility of FA has a crucial effect on binding to FABP and their interaction can be an entire binding mechanism, which allows for recognition of the ligands by the protein.

2 METHODS

2.1 SETTING UP SIMULATIONS

TI and MD simulations where performed using the AMBER suite of programs with the amber ff99SBildn force field (Hornak et al., 2010). The I-FABP structure was taken from the Protein Data Bank. Palmitate and I-FABP complex was created by superimposing the coordinate files of palmitate and I-FABP. The hydrogen bond network of the palmitate and I-FABP complex, including crystallographic water molecules was optimized and hydrogen atoms added with Whatif (Vriend, 1990), and a topology and starting co-ordinate file was created with xleap. The coordinate files of myristate and complex of myristate with I-FABP were created manually by deleting extra carbon and hydrogen atoms of palmitate and then by adding missing hydrogen atoms. A TIP3P water 12 Å truncated octahedron water box and a sodium anion was added to the all coordinate files (Jorgensen et al., 1983).

The created set of coordinates was used in the replicate study (Durrant, 2012). For the current study a different set of coordinates was created by a 100 ps long run of TI simulation (van der Waals creation step) of the myristate conversion into palmitate at 310 K, when λ = 0.6, using soft core potential. The same simulation was done for the complex of myristate and I-FABP ($complex_{myr}$) conversion into the complex of palmitate and I-FABP ($complex_{palm}$). The created systems were used in further TI and MD simulation.

2.2 THERMODYNAMIC INTEGRATION

TI was preformed based on Steinbrecher et al (2007). It consists of three steps. In the first step charge from H28 of myristate was removed, in the second step H28 was changed into–CH_2CH_3 group by adding the van der Waals interactions of the atoms C15, H28, H29, C16, H30, H31, H32, in the third step charge was added on the created atoms.

Each of the steps consists of consecutive minimization, equilibration and production processes using a non-bonded cutoff of 9 Å, i.e. all non-bonded interactions between atoms separated by > 9 Å away are ignored. All steps were performed by using coupling parameters λ from 0.1 till 0.9 with difference 0.1 and λ = 0.025 and 0.975, where λ = 0 has the more myristate like parameters for the ligand, λ = 1 the more palmitate like parameters. TI simulations are run only for the intermediates ($0 < \lambda < 1$) as the way TI is coded in AMBER leads to divide by zero error for λ = 0 and λ = 1.

All three minimization processes were run for 500 cycles starting with 10 cycles of steepest descent minimization followed by conjugate gradient minimization, in constant volume. Equilibration and production processes were run in T = 293 K, using Langevin dynamics with the collision frequency γ = 5 given by gamma_ln (in the case of equilibration) and γ = 2 (in the case of production), pressure was kept constant 1 bar, with pressure relaxation time 0.2 ps. Equilibration processes were run 1 ns long using dt = 0.002 in the first (myristate hydrogen's charge removal) and third step (palmitate additional atoms charge creation) and dt = 0.001 in the second step, while the production process was run as 10 ns, in the case of myristate conversion into palmitate in water, and 14 ns, in the case of $complex_{myr}$ conversion into $complex_{palm}$, following steps 1 ns long using dt = 0.002 in the first and third step and dt = 0.001 in the second step. Convergence of the TI processes was observed by

plotting <dU/dλ> against time for the simulation at each lambda value. <dU/dλ> data was extracted from TI production steps.

In the first and third steps in systems during equilibration and production processes used the SHAKE algorithm, bonds which involved hydrogen atoms were kept constrained, while in the second step systems "shaking" was off. SHAKE usage limits molecular bond stretching, thus it cannot be used in the simulation where atoms are disappearing and appearing. Due to the SHAKE usage equilibration and production processes in step 1 and 3 proceed faster than in step 2.

2.3 FREE ENERGY CALCULATIONS

Calculation of the difference in free energy of binding to I-FABP of palmitate compared to myristate ($\Delta\Delta G_{binding}$) is based on the thermodynamic cycle of the binding of the myristate to I-FABP and of the palmitate binding to I-FABP (Figure 2.1), and it can be explained by the equation $\Delta G_{solution} + \Delta G_P = \Delta G_M + \Delta G_{complex}$, where ΔG_P and ΔG_M are, respectively, the free energy of binding of palmitate and to I-FABP, $\Delta G_{solution}$ and $\Delta G_{complex}$ are, respectively, the free energy for the conversion of myristate into palmitate and $complex_{myr}$ conversion into $complex_{palm}$ (Figure 2.1). Therefore, the difference between free binding energy ($\Delta\Delta G$) of palmitate and myristate binding to I-FABP can be calculated by the equation: $\Delta\Delta G = \Delta G_P - \Delta G_M = \Delta G_{comlex} - \Delta G_{solution}$.

The free energy difference is the integral of $<dU/d\lambda>$ with respect to λ, i.e. the area under the curve, which is determined using the trapezium rule. Firstly, the $<dU/d\lambda>$ values for $\lambda = 0$ and $\lambda = 1$ were extrapolated from the other λ values. The graphs of the $<dU/d\lambda>$ against λ value were prepared for myristate changed to palmitate after 10 ns of simulation time per lambda value and of the $complex_{myr}$ conversion into $complex_{palm}$ after 14 ns of simulation per lambda value. The graphs for the $<dU/d\lambda>$ difference between $complex_{myr}$ conversion into $complex_{palm}$ and myristate conversion into palmitate was determined by subtracting appropriate graphs (see Results).

Calculation of $\Delta\Delta G$ was done in two ways. For one method, firstly area was calculated under the curves of the $<dU/d\lambda>$ vs λ value plot for the conversion of myristate into palmitate and for the conversion of $complex_{myr}$ into $complex_{palm}$. Further, the difference between the area for the conversion of $complex_{myr}$ into $complex_{palm}$ and conversion of solvated myristate into solvated palmitate was calculated by subtraction of these two areas (see Results). In the

second method, the curves from the TI simulations for converting myristate into palmitate in solution and in the I-FABP complex were subtracted prior to integration (see Results).

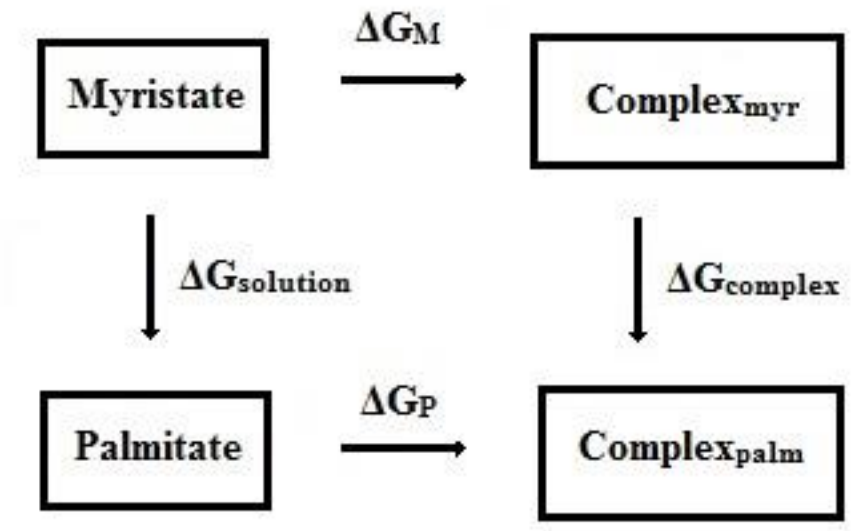


Figure 2.1 Thermodynamic cycle to compute the difference in the free binding energy between myristate and palmitate, while they are binding to I-FABP. $Complex_{palm}$ and $Complex_{myr}$ indicate I-FABP bound respectively to palmitate and myristate, each complex being solvated in water

2.4 MOLECULAR DYNAMICS

MD simulations were performed for myristate in the solution and for the complex of myristate and I-FABP in the solution with a non-bonded cutoff of 9 Å, the temperature was kept at 293 K, with a Langevin thermostat ntt = 3 and collision frequency $\gamma = 5$ given by gamma_ln = 5, pressure was kept constant 1.0 bar with pressure relaxation time 1.0 ps. MD simulations were run for 1ns with dt = 0.002, and were written in output files as 10 consecutive steps. In total a 10 ns long MD simulation was performed. SHAKE was used to keep the positions of hydrogen constrained.

The atomic coordinates were written to a trajectory file every 0.01 ps creating 100 data frames per 1 ns and were used for subsequent analysis of conformations of myristate. In total 1000 data frames containing the information about the atomic coordinates were obtained from 10 ns long MD simulation. Each frame has information about the myristate conformational state.

2.5 CLUSTERING

Myristate clustering was performed by extracting information written in each frame of the trajectory file. Before myristate clustering water molecules and the sodium ion were removed. Two methods for the conformation clustering of free and bound myristate were used – clustering into by the average link algorithm and clustering by dihedral angles, using the program ptraj, supplied with AMBER software.

2.5.1 Clustering by the average link algorithm

Myristate conformations were clustered for co-ordinates from both MD trajectories with 1) just myristate in solution and 2) myristate in the complex with i-FABP, using average-linkage algorithm by comparison of the root mean squared deviation (RMSD) of all atoms myristate molecules were grouped by changing number of clusters from 1 to 10. In both cases myristate clustering was observed by changing MD simulation time from 1ns to 10 ns. Hierarchic arrangement representing the percentage of each cluster was arranged in the figures.

2.5.2 Analysis of the dihedral angles

The value of each distinct dihedral angle of myristate was assigned into one of 36 bins of 10^{o} each, for each recorded snapshot of the simulation. The results were used for calculating the Shannon entropy and also the correlation coefficients comparing the dihedral distribution between protein bound and free myristate.

2.5.2.1 Shannon entropy

The dihedral clustering results were used to calculate the ratio Shannon entropy/the maximum Shannon entropy (H/H_{max}) for each dihedral of free and bound myristate. From the clustering results the probability of each bin for each dihedral was calculated. Each dihedral angle was divided into 36 bins. The maximal entropy is when each bin has equal probability, i.e. p(x) for each bin is 1/36. For each dihedral for each bin -p(x)ln(p(x)) was calculated. The

sum of -p(x)ln(p(x)) for all 36 bin is the Shannon entropy (H) for a dihedral. The maximum Shannon (H_{max}) entropy for one bin is -(1/36)ln(1/36). Thus, H_{max} entropy for a dihedral is -ln(1/36).

2.5.2.2 Correlation coefficient

For each dihedral angle the correlation between the distributions of values observed for the free and bound myristate was calculated. For this purpose data obtained from dihedral angle clustering were used. The Pearson correlation coefficient was calculated using the equation:

$$r_{x,y} = \frac{\sum_{i=1}^{n}(x_i - x_a)(y_i - y_a)}{\sqrt{\sum(x_i - x_a)^2(y_i - y_a)^2}}$$

In the equation x_i and y_i are the number of occurrences in that angle bin for the free and bound fatty acid and x_a and y_a are the mean average occurrence in all angle bins for the free and bound fatty acid.

3 RESULTS

3.1 THERMODYNAMIC INTEGRATION (TI) SIMULATIONS

TI simulations were run for the conversion of sodium myristate into sodium palmitate (here sodium myristate and sodium palmitate will be referred to as myristate and palmitate) in water and when bound to the solvated I-. Three different steps were taken for both conversions. During the first, the charge is removed from the 28^{th} hydrogen of myristate. The second step changes the 28^{th} hydrogen of myristate to the ethyl group of palmitate. The third step adds a charge on the ethyl group.

The time required to perform these simulations is unknown, since the simulation length depends on the system studied and how long it takes to sample the accessible conformations for each intermediate lambda state. Therefore, during the production step for each TI simulation, the change in $\langle dU/d\lambda \rangle$ over time was followed, i.e. $\langle dU/d\lambda \rangle$ needs to converge to a constant value. As it was mentioned before λ value shows the degree of mutation of the sampled ligand. Thus, $\lambda = 0$ has the more myristate like parameters for the ligand, $\lambda = 1$ the more palmitate like parameters. Within the AMBER force field TI is coded to record $\langle dU/d\lambda \rangle$. However, when $\lambda = 0$ and $\lambda = 1$, the analytical expression for $\langle dU/d\lambda \rangle$ in the softcore simulations contains a divided by 0 in its implementation in AMBER and thus gives an error. Therefore, TI simulations are run only for intermediate λ values between 0 and 1.

At first, TI simulation production steps for myristate conversion into palmitate and $complex_{myr}$ conversion into $complex_{palm}$ were run as 10 consecutive simulations, each 1ns long. Data from TI simulations, where myristate was converted to palmitate show that the processes have been run long enough, all curves are observed as almost straight lines (Figure

3.1). For the TI simulations, where the charge was removed from the 28th hydrogen of myristate, there were no high changes observed in <dU/dλ> for all λ values (Figure 3.1), which indicates likely convergence.

At the beginning of the TI simulations for myristate conversion into palmitate, when the 28th hydrogen of myristate was converted into the ethyl group, for some λ values there was a fluctuation in <dU/dλ> (Figure 3.1). The most notable change in <dU/dλ> was observed for λ = 0.8. Nevertheless, for all expected convergence is observed after 5 ns.

For TI simulations when a charge was added on the ethyl group, after 2 ns there was observed a small fluctuation of <dU/dλ> was observed for some λ values (Figure 3.1). However, similar to the TI simulation, when the charge was removed from the 28th hydrogen, all curves quickly become almost straight lines. Thus, the TI simulations for the conversion of myristate into palmitate in water indicate appropriate convergence within 6 ns or less.

For TI simulations where the charge was removed from the 28th hydrogen of myristate for the conversion of complex$_{myr}$ into complex$_{palm}$ high fluctuations of <dU/dλ> were observed for all λ values for the first 10 ns (Figure 3.2). Therefore, TI production steps for all λ values were continued for a further 4 ns. When TI simulations were continued for the next 4 ns, a stabilization of <dU/dλ> was observed. Moreover, for λ values except λ = 0.1 <dU/dλ> are almost constant after 12 ns of TI simulations. For λ = 0.1 <dU/dλ> is constant from 9 to 12 ns. When the TI simulation for λ = 0.1 was run for >12 ns another increase in <dU/dλ> was observed (Figure 3.2 (a)), suggesting that this simulation should be continued for a longer time.

In the second step of TI simulations for complex$_{myr}$ conversion into complex$_{palm}$, when the hydrogen of myristate was mutated into the additional ethyl group found in palmitate, the fluctuation of <dU/dλ> is much higher than was observed in the first production step. A small

<dU/dλ> fluctuation was observed for λ = 0.4 and 0.9, but for λ = 0.1, 0.8 and 0.975, <dU/dλ> remains almost constant (Figure 3.2 (b)) after 10 ns of TI simulations. While some of the curves observed in Figure 3.2 (b) show that after 10 ns of TI simulations the fluctuation of <dU/dλ> for λ = 0.025, 0.2, 0.3, 0.5, 0.6 and 0.7, is relatively high. Therefore, TI production steps for all λ values were continued for a further 4 ns. When TI simulations were run for a further 4 ns <dU/dλ> became almost constant for most values of λ. Nevertheless, for λ = 0.2, there was observed an increase in <dU/dλ> during the last ns. The curve for λ = 0.3, which was observed as a constant from 5 to 9 ns, shows a decrease in <dU/dλ> from 11 to 13 ns (Figure 3.2 (b)). Therefore, to verify that, in further simulations, there is no significant change in <dU/dλ> and to stabilize the change in <dU/dλ> for the processes where λ = 0.2 and 0.3, all TI simulations should be run for at least 5 ns longer.

In the third step of TI simulations of $complex_{myr}$ conversion to $complex_{palm}$ when a charge was added to the ethyl group, fluctuations in <dU/dλ> were seen for λ = 0.025, 0.3, 0.5, 0.7 and 0.8 for the first 10 ns, but <dU/dλ> converged to a stable value for λ = 0.4, 0.9 and 0.975 from 8 to 10 ns and for λ = 0.1 and 0.2 from 6 to 10 ns (Figure 3.2 (c)). After a further 4 ns of TI simulations the almost constant lines observed for λ = 0.1 and 0.2 indicated a convergence in the value of <dU/dλ>, which was also seen from 11ns for λ = 0.025 and from 12 ns for λ = 0.5. Nevertheless, <dU/dλ> for λ = 0.4, 0.7, 0.9 and 0.975, which became almost constant by 10 ns, slightly decreased when TI was run for a longer time.

A TI production step of 1 ns for a simulation of a ligand free in solution takes about 30 hours, but for ligand and protein complex 40-60 hours of computational time. Within the time limitations of this project it was not possible to continue the TI simulation for longer and thus the results here were used for the calculation of free binding energy, although it is desirable to run the simulations for longer.

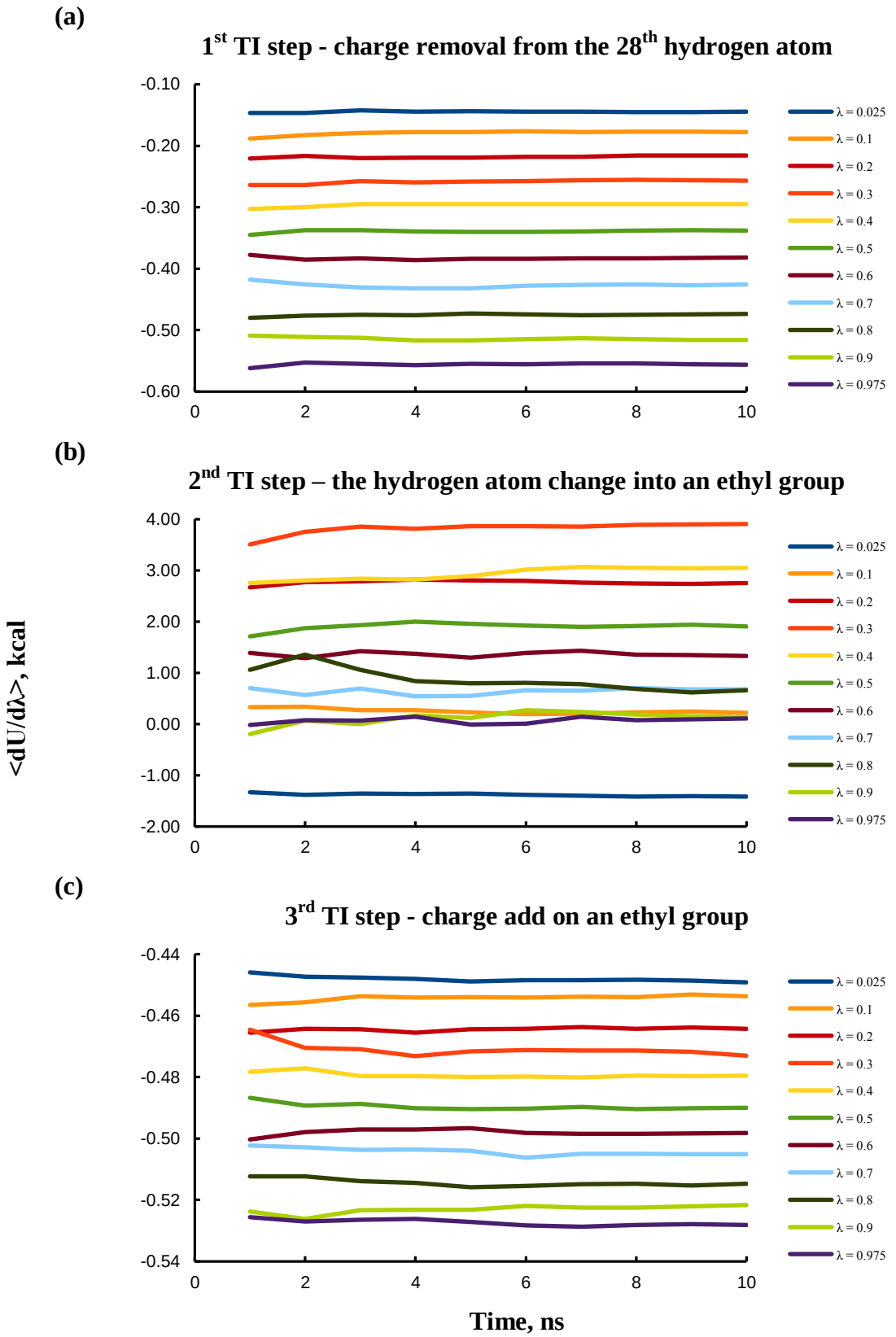


Figure 3.1 The change in <dU/dλ> over time for the conversion of sodium myristate to sodium palmitate in water solution. (a) 1st TI step - charge removal from the 28th hydrogen atom, (b) 2nd TI step – the hydrogen atom is changed into an ethyl group, (c) 3rd TI step - charge add on the ethyl group.

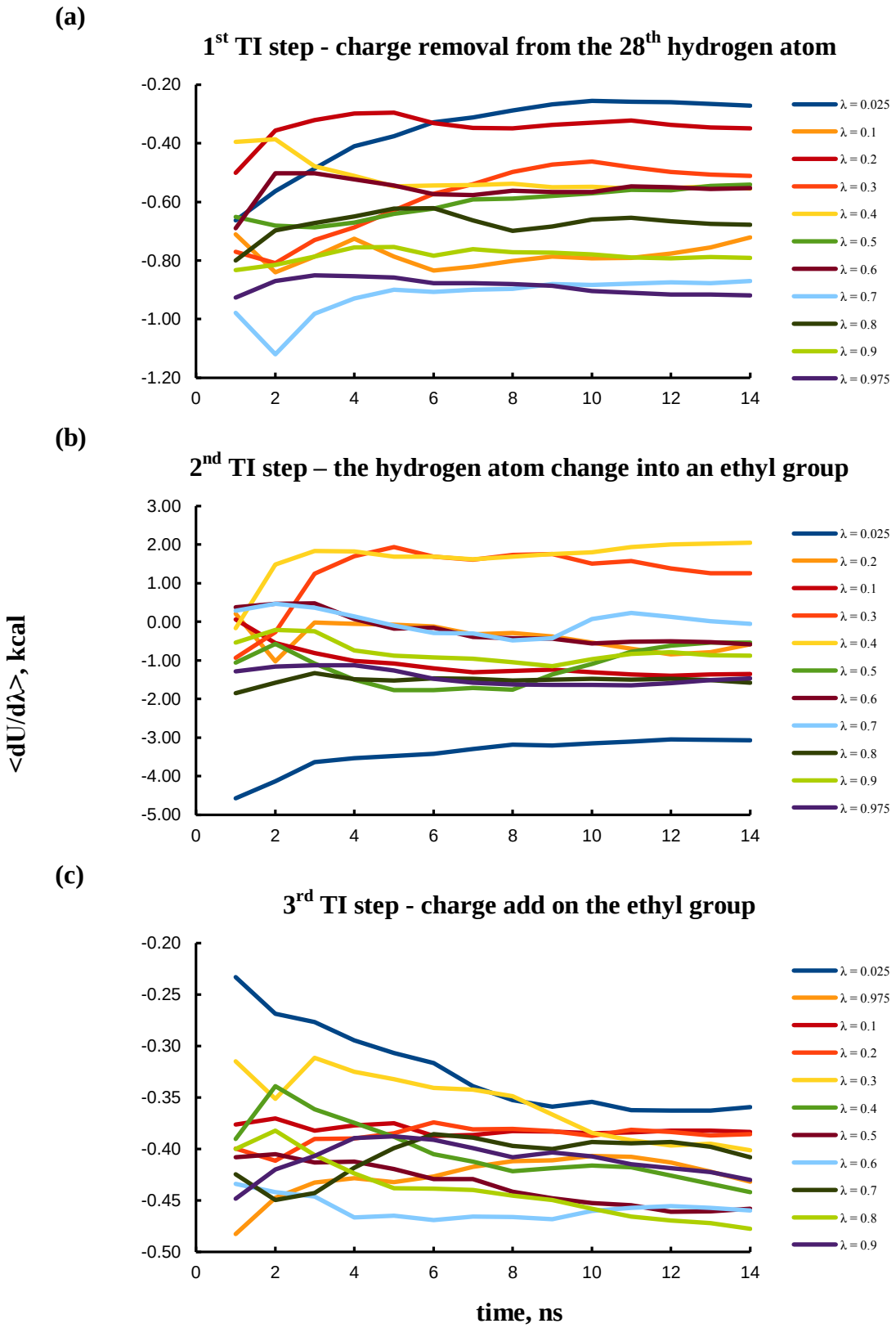


Figure 3.2 The change of <dU/dλ> over time for complex$_{myr}$ conversion to complex$_{palm}$ in water. (a) 1st TI step - charge removal from the 28th hydrogen atom, (b) 2nd TI step – the hydrogen atom is changed into an ethyl group, (c) 3rd TI step - charge add on the ethyl group.

3.2 CALCULATION OF THE DIFFERENCE BETWEEN THE FREE ENERGY OF MYRISTATE AND PALMITATE BIND TO I-FABP

In the Introduction section it was explained that from TI simulations can be calculated only the difference of free binding energy (ΔΔG) between two very similar ligands. To calculate ΔΔG, graphs of <dU/dλ> with respect to λ were plotted. As mention before, TI simulations cannot be run for λ = 0, and 1. Therefore, <dU/dλ> values for λ = 0 and 1 were extrapolated from the other λ values. For myristate converted to palmitate in solution, the graphs were plotted after 10 ns of TI simulations (Figure 3.3 (a)). In the case of $complex_{myr}$ conversion into $complex_{palm}$ the graphs were prepared after 14 ns of simulation (Figure 3.3 (b)). The difference between <dU/dλ> for these two process, i.e. between the conversion of $complex_{myr}$ into $complex_{palm}$ and the conversion of myristate into palmitate are plotted against lambda in figure 3.3 (c) and were calculated by simply subtracting the lines in 3.3 (a) from 3.3 (b), e.g. figure 3.3 (c) column 1 is a subtraction of figure 3.3 (a) column 1 from figure 3.3 (b) column 1.

The change in <dU/dλ> with respect to λ for the TI simulations where myristate was converted into palmitate is smooth (Figure 3.3 (a)). A relatively smooth change in <dU/dλ> with respect to λ was also observed in the case of $complex_{myr}$ changing into $complex_{palm}$, where a charge was added to the ethyl group (Figure 3.3 (b) column 3). For the removal of the charge from the 28^{th} hydrogen of the myristate the change of <dU/dλ> with λ is smooth, except λ = 0.1 and 0.7, which deviate from the general trend of the curves (Figure 3.3 (b) column 1). For the conversion of $complex_{myr}$ into $complex_{palm}$, when the 28^{th} hydrogen was changed into an ethyl group the change in <dU/dλ> with respect to λ does not form a smooth curve (Figure 3.3 (b) column 2). The lack of smooth <dU/dλ> versus λ curves could arise due to insufficient conformational sampling at a given λ point, i.e. the point is erroneous or from

not having enough λ points. Either way it is likely to contribute to errors in the calculation, and continuation of this aspect the project should investigate this.

Comparing graphs plotted for the conversion of myristate into palmitate and $complex_{myr}$ into $complex_{palm}$ the graphs for similar production steps of the change in <dU/dλ> with respect to λ are mapping similar curves. For example, for the charge removal from the 28^{th} hydrogen of myristate, curves for the conversion of myristate into palmitate and $complex_{myr}$ into $complex_{palm}$ both have a negative gradient. Therefore, the graphs for the subtraction of <dU/dλ> in figure 3.3 (a) from the <dU/dλ> in figure 3.3 (b) are expected to form straight lines with slight or no slope. Although, the tendency to form these straight lines is observed in plotted graphs (Figure 3.3 (c)), fluctuation of the change in <dU/dλ> with respect to λ is observed.

Both graphs of <dU/dλ> over time and graphs of <dU/dλ> with respect to λ show that TI simulations for $complex_{myr}$ conversion into $complex_{palm}$ were not run long enough. Therefore, two ΔΔG values were calculated, $\Delta\Delta G^{*}$ was calculated from TI simulations of myristate conversion into palmitate and from $complex_{myr}$ conversion to $complex_{palm}$ run for 10 ns, and $\Delta\Delta G^{**}$ was calculated, when TI simulations for the conversion of myristate into palmitate was run for 10ns, but TI simulation of $complex_{myr}$ conversion to $complex_{palm}$ for 14 ns. Comparison of these values should give some indication of the likely effects on ΔΔG arising from possibly not having been able to run the simulations for long enough.

The ΔΔG calculation is based on integration of <dU/dλ> with respect to λ. An integral of <dU/dλ> is determined by calculation of the area under curves in the prepared graphs of <dU/dλ> with respect to λ. The area under the curves, thus integral of <dU/dλ> is determined using the trapezium rule.

Calculation of ΔΔG was done in two ways. For one method, firstly by calculating the areas (A) under the curves of <dU/dλ> with respect to λ for myristate change into palmitate (Figure 3.3 (a)) and under the curves of <dU/dλ> with respect to λ for $complex_{myr}$ conversion to $complex_{palm}$ (Figure 3.3 (b)) for all three TI steps. Next, the difference between the areas (ΔA) of $complex_{myr}$ conversion to $complex_{palm}$ and myristate conversion to palmitate by subtraction was calculated, e.g. ΔA_1 (the difference between the areas for the first TI step) is a subtraction of A (Figure 3.3 (a) column 1) from A (Figure 3.3 (b) column 1). In the same way a calculation was made for the difference in area for the second (ΔA_2) and third (ΔA_3) TI simulation step. ΔΔG is a sum of the ΔA_1, ΔA_2 and ΔA_3 (Figure 3.4 (a)).

The second ΔΔG calculation used only one integration step since the curves for simulations in solution and in the I-FABP complex were subtracted prior to integration, e.g. $<dU/d\lambda_i>$ (Figure 3.3 (c)) = $<dU/d\lambda_i>$ (Figure 3.3 (b)) - $<dU/d\lambda_i>$ (Figure 3.3 (a)) for λ_i.

Our calculated ΔΔG shows that ΔG_M > than ΔG_P, i.e. palmitate binds the strongest. To verify the obtained ΔΔG and how accurate is our simulation system, ΔΔG was calculated from experimentally obtained data. Velkov et al. (2005 and 2007) obtained dissociation constants (K_d) of palmitate (K_{dP}) and myristate (K_{dM}) during 1-anilinonaphthalene-8-sulfonic acid (ANS) displacement by the ligands in the complex of ANS and I-FABP. Velkov et al. (2005 and 2007) performed experiments using three different buffers. The difference of the palmitate and myristate binding energy (ΔΔG) was calculated using equation:

$$\Delta\Delta G = \Delta G_P - \Delta G_M = RTlnK_{dP} - RTlnK_{dM} = RTln(K_{dP}/K_{dM})$$

In the equation $K_d = [L][P]/[PL]$, where [L] is the concentration of the ligand, [P] is the concentration of the protein and [PL] is the concentration of the complex. $\Delta G_{dissociation} = -RTlnK_d$ for the reaction $PL \rightarrow P + L$. The $\Delta G_{binding}$ goes in the opposite direction so $\Delta G_{binding} = RTlnK_d$; $\Delta G_P = RTlnK_{dP}$; $\Delta G_M = RTlnK_{dM}$. Calculated ΔΔG from experimental data shows

that ΔΔG is dependent on the buffer conditions used. Nevertheless, the calculated ΔΔG are similar (Figure 3.4 (c)).Comparing our obtained ΔΔG with the ΔΔG calculated from experimental data it is clear that our results overestimate the magnitude of the ΔΔG value, although the simulations performed here are in "pure water" compared to the different buffer conditions used in the experiments and are thus not directly comparable. Nevertheless, ΔΔG calculated from experimental data as well as our calculated ΔΔG show that ΔG_M > than ΔG_P, i.e. palmitate binds the strongest.

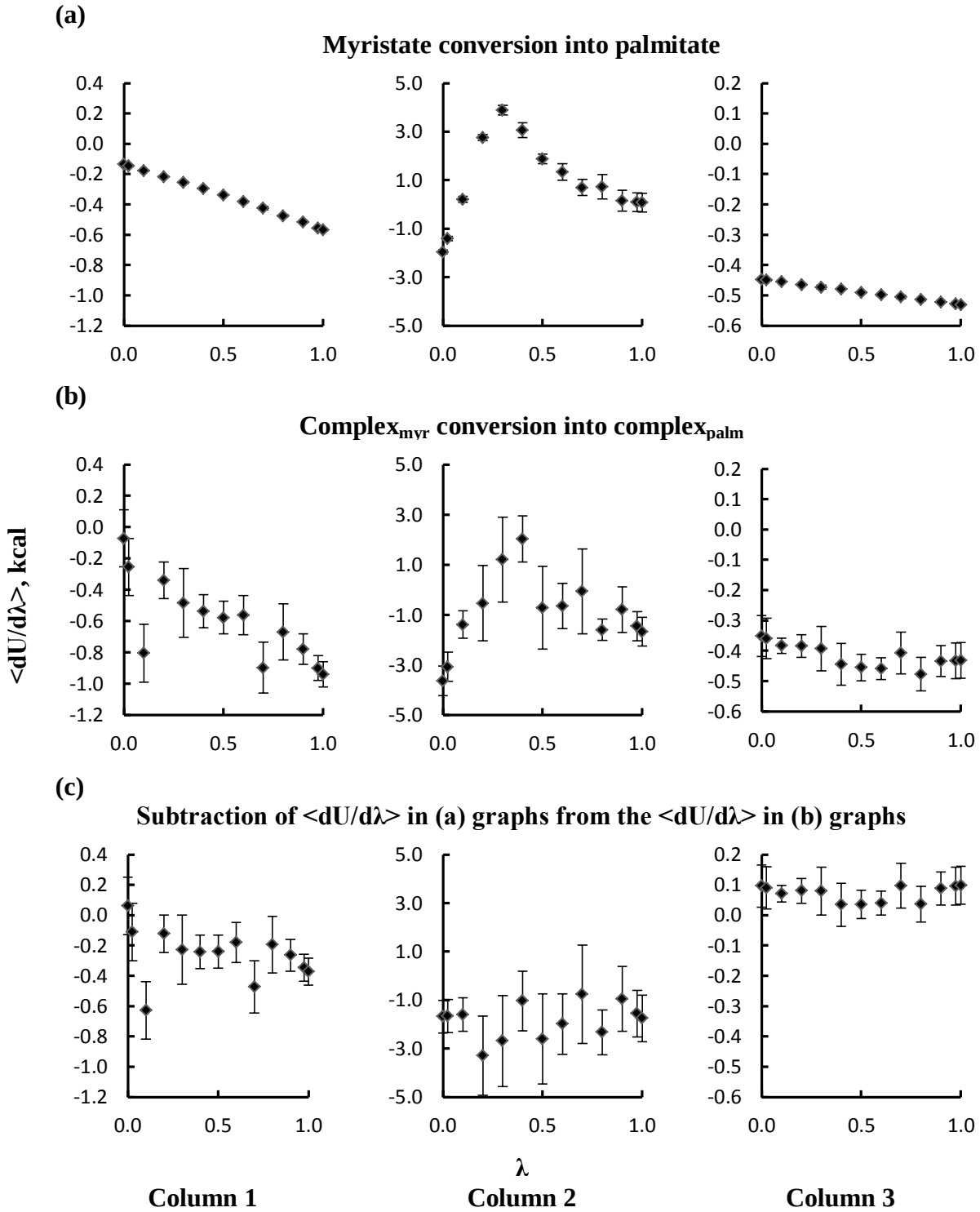


Figure 3.3 The change of <dU/dλ> with respect to λ (a) TI simulations for the conversion of myristate into palmitate, (b) TI simulations for the conversion of complex$_{myr}$ into complex$_{palm}$, (c) graphs of <dU/dλ> with respect to λ calculated by subtraction of <dU/dλ> in (a) graphs from <dU/dλ> in (b) graphs. Column 1 is 1st TI step - charge removal from the 28th hydrogen atom, column 2 is 2nd TI step - the hydrogen atom change into an ethyl group and column 3 is 3rd TI step - charge was added onto the ethyl group. In each graph λ = 0 has the more myristate like parameters for the ligand, λ = 1 the more palmitate like parameters.

(a)

Obtained ΔΔG from TI simulations

Method	$\Delta\Delta G^{*}$, kJ/mol	$\Delta\Delta G^{**}$, kJ/mol
1.	-9.080	-8.816
2.	-7.573	-8.795

(b)

Calculated ΔΔG from reported experimental K_d data

No	Buffer	ΔΔG, kJ/mol
1.	50 mM Tris-HCl, 150 mM NaCl, 1 ml dithiothreitol, 0.5 mM EDTA	-1.747
2.	50 mM Tris-HCl, 150 mM NaCl	-1.220
3.	20 mM Tris-HCl	-1.380

Figure 3.4 The difference of the myristate and palmitate free binding energy (ΔΔG), while they bind to I-FABP. (a) obtained ΔΔG from TI simulations, where $\Delta\Delta G^{*}$ is calculated from TI simulations run for 10 ns, $\Delta\Delta G^{**}$ is calculated from TI simulations of myristate conversion to palmitate run for 10 ns and from TI simulations of $complex_{myr}$ conversion to $complex_{palm}$, run for 14 ns (b) Calculated ΔΔG from reported experimental K_d data using different buffers. K_d for No 1 ΔΔG calculation is taken from Velkov et al. (2005). While for K_d for No 2 and 3 ΔΔG calculation is taken from Velkov et al. (2007).

3.3 INVESTIGATING THE FLEXIBILITY OF FREE MYRISTATE IN SOLUTION AND OF BOUND MYRISTATE TO I-FABP

MD simulations of myristate in water and $complex_{myr}$ in water were run as 10 consecutive steps, each 1ns long with 1000 frames (co-ordinate sets) written to file. Each frame has information about the myristate conformational state. Before clustering the myristate co-ordinate sets water molecules and the sodium ion was removed. Two methods for conformational analysis of free and I-FABP bound myristate data were used – clustering by the average link algorithm and classification by the value of the dihedral angles. Clustering by the average link algorithm was performed by atom distance comparison using RMSD. Classification by dihedral angles was performed to for each dihedral angle, allocating each dihedral angle to a given bin for each given simulation frame, dependent upon the value of the dihedral angles.

Myristate consists of 14 carbon atoms as represented in figure 3.5. Thus, the dihedral angle for the 1st dihedral is angle between bond O- C_1 and C_2-C_3 in a chain O-C_1-C_2-C_3.

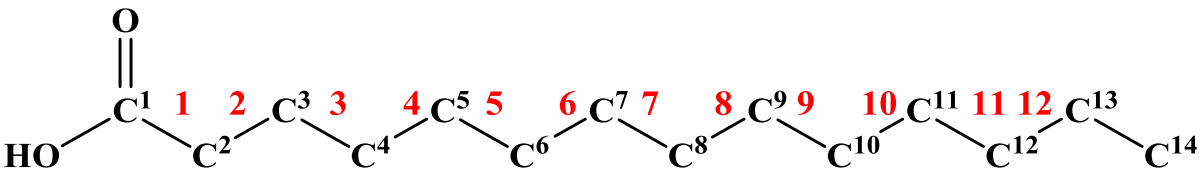


Figure 3.5 Formula of myristate, with carbons and dihedral angles labeled the latter in red.

3.3.1 Clustering by the average link algorithm

Conformations of free myristate and bound myristate to I-FABP were grouped into clusters from one to five according to their structural similarities. Clustering data is presented in figure 3.6.

Free myristate has one dominant conformation, when conformations are grouped into two to five clusters. Grouping myristate conformation into five clusters, the dominant cluster consists of 65.2%, whilst the second largest of 29.3% from the total count of myristate conformations. Moreover, when myristate conformations were grouped into three clusters these two clusters were grouped as one. Therefore, the biggest two clusters, when myristate conformations are grouped into five clusters, consist of similar myristate conformations (Figure 3.6 (a)). The average RMSD distance of all cluster members from the average structure of the cluster is set by default as < 0.6 Å indicating that each representative cluster is compact.

Grouping free myristate conformations into two clusters, the dominant cluster consists of 96.5%. Clusters from simulations with myristate bound to I-FABP consist of 54.2% and 45.8% from the total count of myristate conformations. When the number of clusters grouping conformations of bound myristate was increased from two to five the percentage of the biggest cluster decreases from 54.2% to 32.5%, while the percentage of the second biggest cluster stays unchanged – 45.8% (Figure 3.6 (b)).

Representative structures of clusters, when conformations of free and bound myristate are grouped into five clusters, were observed using pymol (Figure 3.7). For both free and bound myristate comparing cluster representative structures by their dihedral angles, the high similarity between the two biggest clusters was confirmed. The difference in the angle higher than 5° was observed at the 1^{st}, 2^{nd}, 6^{th} and 7^{th} dihedrals for free myristate - 60^{o}, 90^{o}, 100^{o} and

100°, respectively (Figure 3.7 (a)). For bound myristate the difference in angle higher than 5° was observed for 3rd, 4th, 6th and 7th dihedrals, it is about 60°, 30°, 90° and 100°, respectively (Figure 3.7 (b) and (c)).

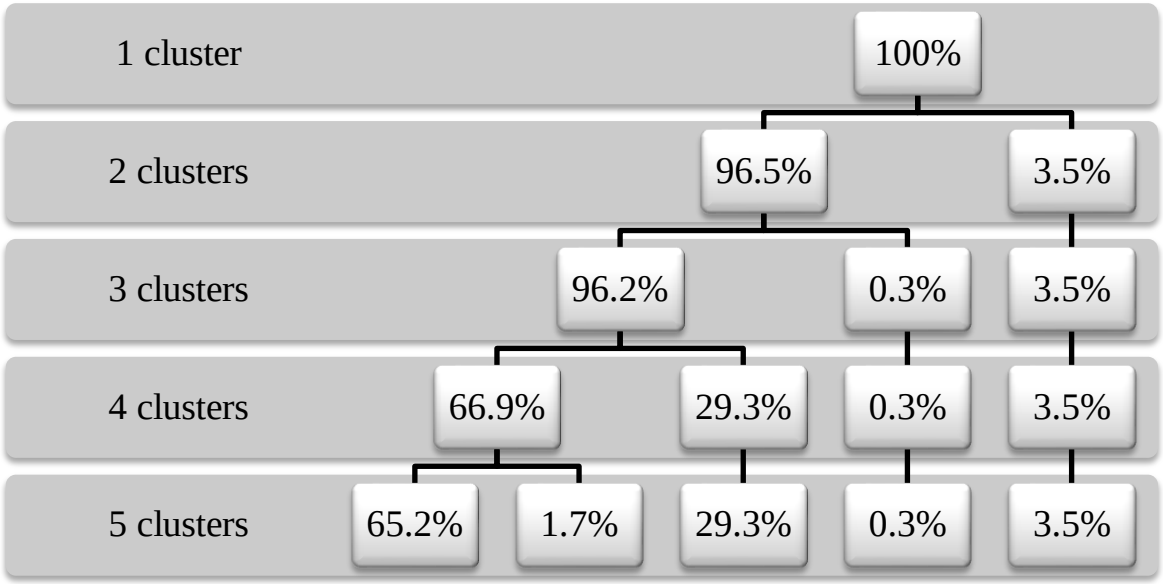
Comparing cluster representative structures for the biggest clusters of free and bound myristate with the third cluster, respectively, six and seven dihedrals from the total of twelve were found to form a similar angle (maximal angle difference is within 5°). Five dihedrals from the total of twelve were found to form a similar angle comparing the biggest cluster of free myristate with either the fourth or the fifth cluster of free myristate. In the case of bound myristate six dihedrals were found to form a similar angle comparing the biggest cluster with either the fourth or the fifth cluster.

. Therefore, clusters of bound myristate look like being more similar, than free myristate clusters are. However, the sum of conformation of two biggest clusters for free and bound myristate is 945 and 783 from total 1000 myristate conformations.

An important observation was that, when the conformations are grouped into five clusters, the biggest clusters of free and bound myristate have extremely high structural similarity (Figure 3.8), i.e. the dominant conformation myristate found in the I_FABP bound structure is almost identical to the dominant conformation found for myristate free in solution. A difference was observed only at the 1st and 2nd dihedral angles.

(a)

Free myristate clusters



(b)

Bound myristate clusters

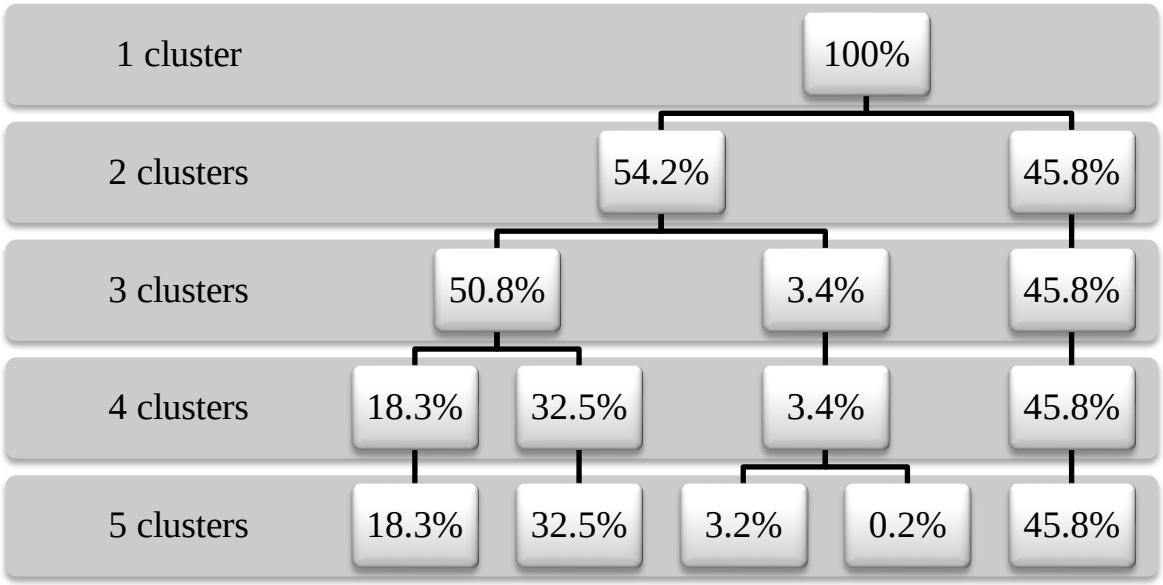


Figure 3.6 The clustering of myristate into distinct clusters. The clusters were divided from the linkage algorithm based on RMSD between structures from 1000 snapshots saved from a 10 ns simulation, with the algorithm being set to return from one to five clusters. (a) The diagram represents the cluster connection and the percentage of free myristate conformations in each cluster. (b) The diagram represents the cluster connection and the percentage of bound myristate conformations in each cluster.

(a)

Myristate in solution

(b)

Myristate bound to I-FABP (front view)

(c)

Myristate bound to I-FABP (90º rotation $\overrightarrow{\circlearrowright}$X)

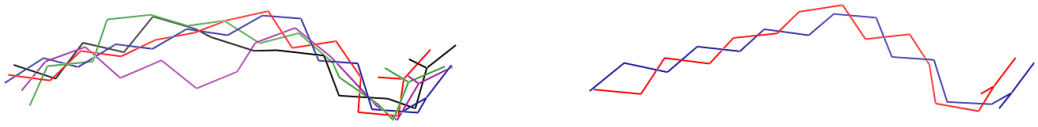


Figure 3.7 Representative structures for five clusters that represent the dominant conformations in simulations of myristate free in solution and bound to I-FABP. The clusters were derived from the linkage algorithm based on RMSD between structures from 1000 snapshots saved from 10 ns simulations. (a) Conformational clustering of free myristate in solution. Myristate is represented in green, red, black, blue and purple respectively consist of 65.2%, 29.3%, 3.5%, 1.7% and 0.3% recorded myristate conformations. (b) Conformational clustering of bound myristate to I-FABP (front view). Myristate is represented in blue, red, green, purple and black respectively consist of 45.8%, 32.5%, 18.3%, 3.2% and 0.2% recorded myristate conformations. (c) Conformational clustering of bound myristate to I-FABP (90º rotation). Myristate is represented in blue, red, green, purple and black respectively consist of 45.8%, 32.5%, 18.3%, 3.2% and 0.2% recorded myristate conformations.

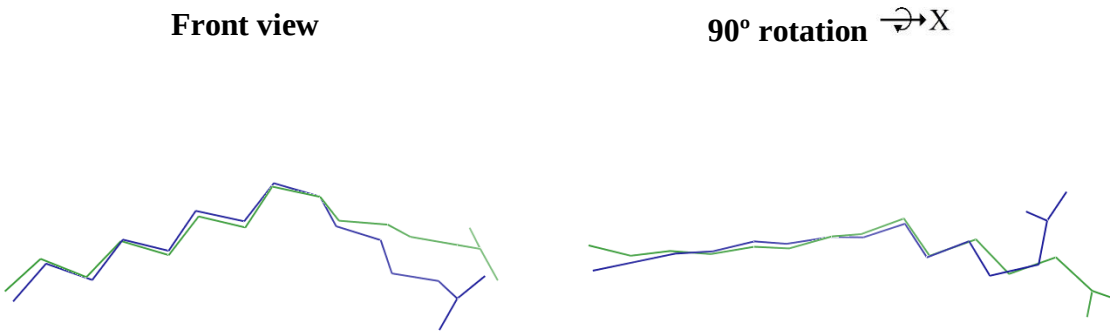


Figure 3.8 Comparison of the representative conformations for the dominant cluster for free and bound myristate. Free myristate is represented in green, bound myristate in blue. Five clusters were derived from the linkage algorithm based on RMSD between structures from 1000 snapshots saved from 10 ns simulations.

.

3.3.2 Clustering by dihedral angles

Clustering by dihedral angles was performed for free myristate in solution and for bound myristate for each dihedral of myristate, using 36 bins. The distribution of the formed angles for each dihedral is represented in graphs (Figure 3.9 and 3.10).

The angle distribution at the 1st dihedral for free myristate was observed as four preferred angle ranges centered at 80º, 90º, -120º and -80º. Where first two and last two overlap forming two broad bands. At the 2nd to 12th dihedrals free myristate has three preferred angle ranges centered at 60º, 175º and -70º. The dominantly preferred angle is centered at 175º.

The angle distribution at the 1st dihedral for bound myristate was observed as two preferred angle ranges. The dominantly preferred angle range is centered at -120º, the second preferred angle at -70º. The angle distribution at the 2nd, 3rd and 5th dihedrals also was observed as two preferred angle ranges. The dominantly preferred angle and the second preferred angle are centered at 55º and -70º for the 2nd dihedral, at 175º and -100º for the 3rd dihedral and at 175º and -70º for the 5th dihedral, respectively. At the 4th and 6th to 12th dihedrals three preferred angle ranges centered at 65º, 175º and -70º was observed, where the dominantly preferred angle is centered at 175º.

Shannon entropy divided by the maximum possible Shannon entropy (H/H_{max}) was calculated for each dihedral from the data shown in figures 3.9 and 3.10. H/H_{max} values for each dihedral are shown in figure 3.11 (a). $H/H_{max} = 1$ refers to the maximal dihedral flexibility, while $H/H_{max} = 0$ to the maximal rigidity. Dihedrals of free myristate were arranged in order with increasing in flexibility: $5^{th} \approx 9^{th} < 6^{th} = 10^{th} \approx 7^{th} = 3^{rd} \approx 8^{th} \approx 11^{th} \approx 4^{th} = 12^{th} \approx 2^{nd} < 1^{st}$, where H/H_{max} increases from 0.60 to 0.88. While dihedrals of bound

myristate were arranged in order with increasing in flexibility: $10^{th} < 8^{th} \approx 5^{th} = 2^{nd} = 12^{th} \approx 9^{th} = 1^{st} \approx 3^{rd} < 11^{th} < 7^{th} < 4^{th} < 6^{th}$. H/H_{max} for bound myristate increases from 0.54 to 0.76.

Comparing H/H_{max} values at the same dihedral between free and bound myristate was confirmed that myristate became more rigid, when it binds to I-FABP. Dihedrals of myristate were arranged in order of the decreasing change of flexibility: $11^{th} \approx 3^{rd} < 10^{th} = 8^{th} \approx 12^{th} \approx 2^{nd} < 1^{st}$. It is intriguing that myristate became more flexible at the 6^{th} dihedral, when it binds to I-FABP. No change in flexibility at the 4^{th}, 5^{th}, 7^{th} and 9^{th} dihedrals was observed as the change in $H/H_{max} \leq 0.3$.

The correlation coefficient of each dihedral angle was calculated to compare the angle distribution between free and bound myristate (Figure 3.11 (b)). The correlation coefficient = 1 refers to the identical angle distribution. Dihedrals were arranged in order with increasing similarity between free and bound myristate angle distribution: $2^{nd} < 1^{st} < 6^{th} < 4^{th} < 3^{rd} \approx 7^{th} = 12^{th} \approx 5^{th} = 8^{th} \approx 9^{th} = 11^{th} \approx 10^{th}$. The smallest correlation coefficient < 0.25 was observed at the 1^{st} and 2^{nd} dihedrals. Thus, at this dihedrals angle distribution difference is the highest. The correlation coefficient for the 6^{th} dihedral is 0.88, for the 3^{rd} to 5^{th} and 7^{th} to 12^{th} is > 0.90, which indicates that free and bound myristate have very similar angle distribution at these dihedrals.

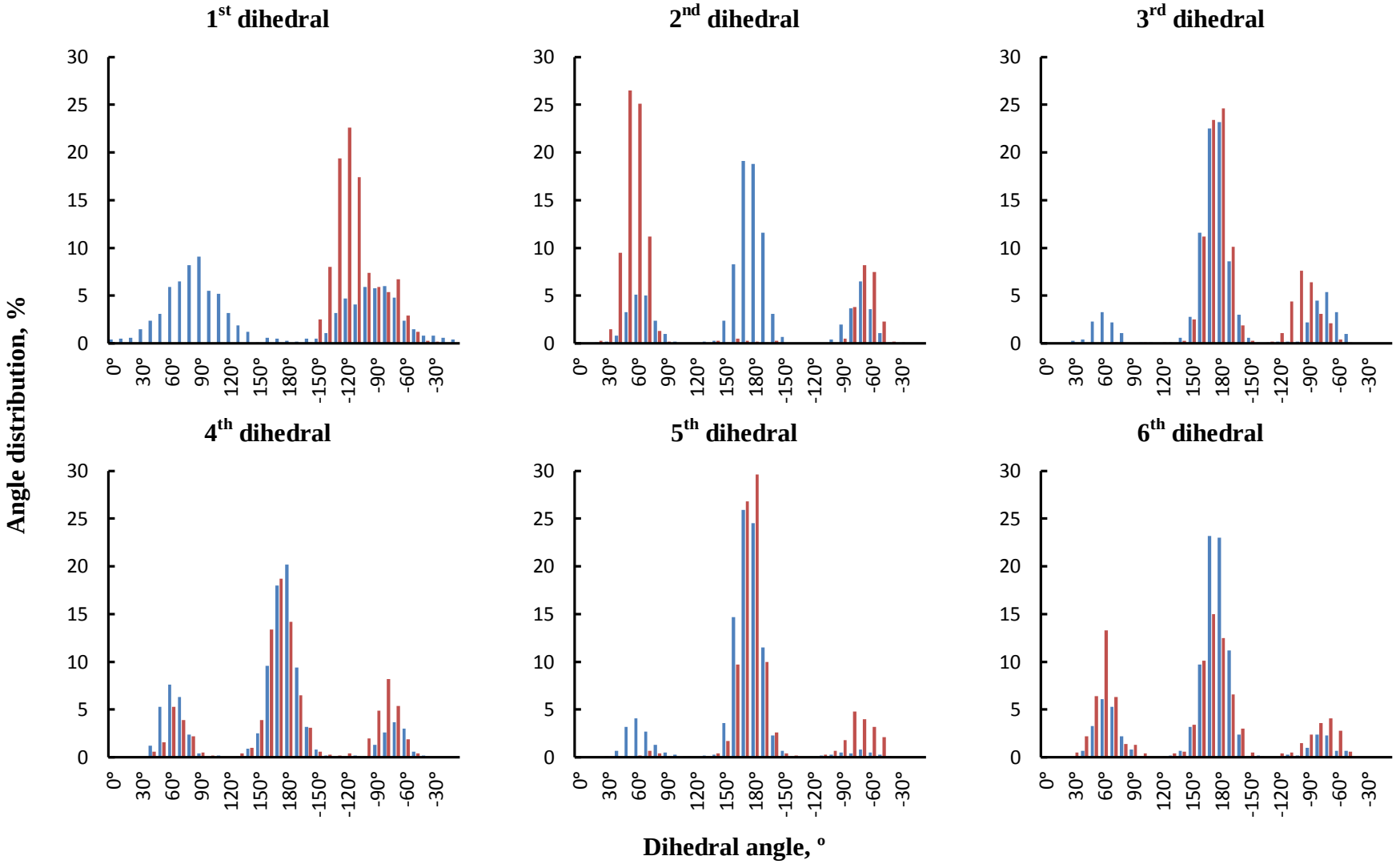


Figure 3.9 Myristate clustering by dihedral angles using 36 bins of 10º (dihedrals 1st to 6th). The angle distribution at dihedrals for free myristate is presented in blue, for bound myristate in red.

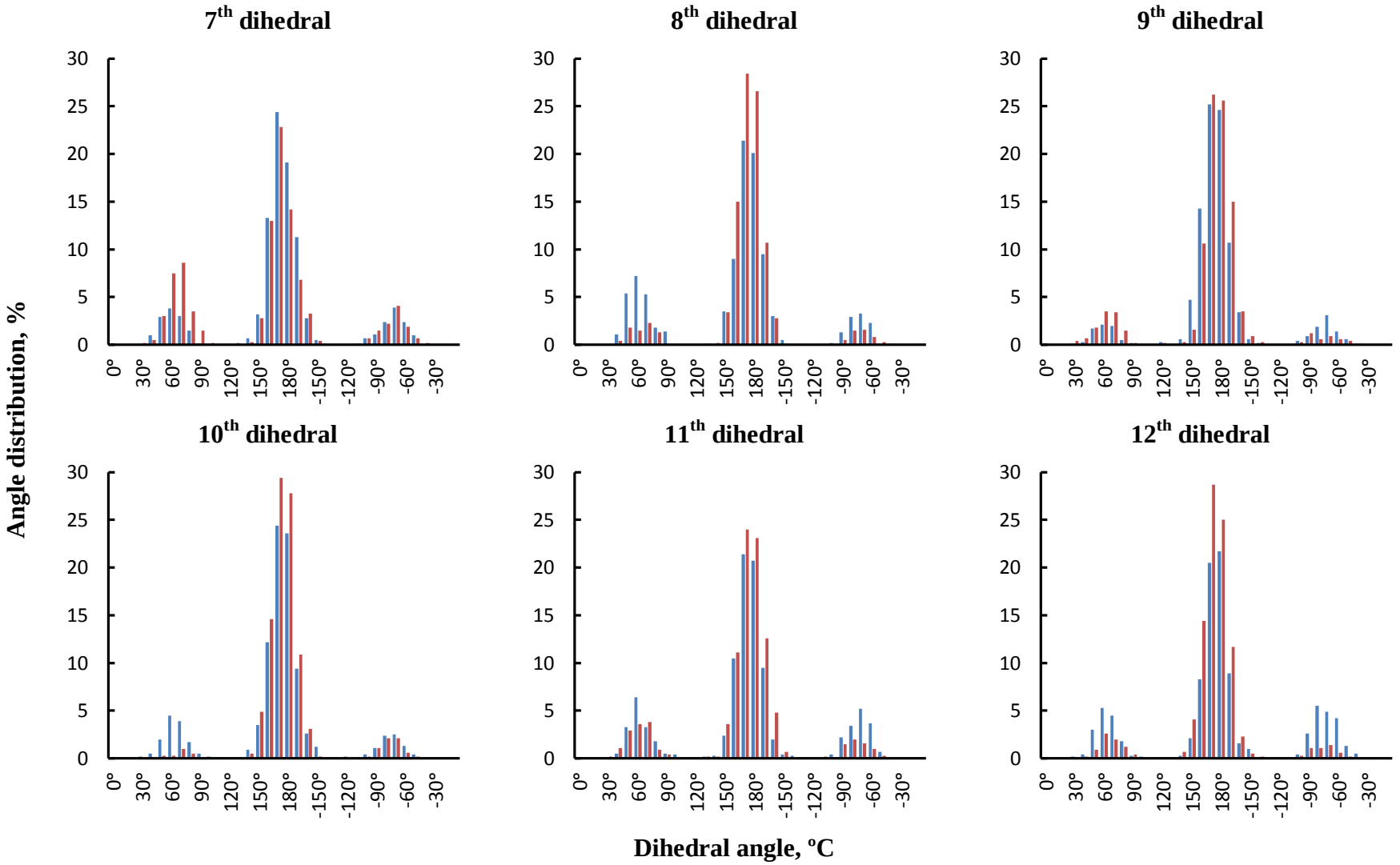


Figure 3.10 Myristate clustering by dihedral angles using 36 bins of 10° (dihedrals 7th to 12th). The angle distribution at dihedrals for free myristate is presented in blue, for bound myristate in red.

(a)

Dihedral number	H/H_{max}			
	Myristate in solution	Myristate bound to I-FABP	Palmitate in solution	Palmitate bound to I-FABP
1st	0.88	0.60	0.89	0.47
2nd	0.71	0.58	0.73	0.40
3rd	0.67	0.61	0.61	0.49
4th	0.70	0.73	0.67	0.43
5th	0.60	0.58	0.58	0.60
6th	0.65	0.76	0.58	0.49
7th	0.67	0.70	0.71	0.59
8th	0.68	0.57	0.65	0.64
9th	0.61	0.60	0.67	0.51
10th	0.65	0.54	0.71	0.70
11th	0.69	0.64	0.65	0.51
12th	0.70	0.58	0.62	0.61
13th	-	-	0.67	0.63
14th	-	-	0.65	0.73

(b)

Dihedral number	Correlation coefficient	
	Myristate	Palmitate
1st	0.24	0.55
2nd	0.11	0.01
3rd	0.94	0.96
4th	0.91	0.90
5th	0.96	0.05
6th	0.88	0.98
7th	0.95	0.85
8th	0.96	0.96
9th	0.98	0.92
10th	0.99	0.77
11th	0.98	0.98
12th	0.95	0.98
13th	-	0.81
14th	-	0.97

Figure 3.11 The determination of the change in myristate and palmitate flexibility, when myristate and palmitate binds to I-FABP. (a) Shannon entropy divided by the maximal possible Shannon entropy (H/H_{max}), where $H/H_{max} = 1$ refers to the maximal possible flexibility and $H/H_{max} = 0$ to the minimal possible flexibility (maximal rigidity). (b) The correlation coefficient for the angle distribution each dihedral between free and bound myristate and palmitate (H/H_{max} and correlation coefficient values for palmitate are taken from Durrant, 2012).

4 DISCUSSION

4.1 CONVERGENCE OF TI SIMULATIONS

The graphs recording the change in <dU/dλ> over time for the conversion of myristate into palmitate show that the processes for all λ and all production steps have been run long enough, all curves are observed as almost straight lines.

In the case of the conversion of $complex_{myr}$ into $complex_{palm}$, convergence was observed only for some λ values after being run for 14 ns. However, for most λ values $complex_{myr}$ has higher energy changes and needs to be simulated for a longer time for the energy to converge to a stable value compared to the free myristate simulation. Slow convergence of ensemble averages and high energy changes indicate greater conformational changes within the system and highlight the importance of protein flexibility. Within the time limitations of this project it was not possible to continue the TI simulation for longer and thus the results here were used for the calculation of free binding energy, although it is desirable to run the simulations for longer.

The convergence of our TI simulations was compared with a replicate study (Durrant 2012). Most conditions of TI simulations for both projects were set up as the same. The difference between the projects is set up of pressure relaxation time 0.2 ns and 1 ns, in this project and Durrant's project, respectively.

Expected convergence of TI simulation for myristate conversion into palmitate was observed only for production step 1 and step 3 after 10 ns (Durrant 2012). Despite observed convergence, the fluctuation of <dU/dλ> over time is higher than it was observed in our TI simulation for the same steps. Moreover, for the conversion of myristate into palmitate when the 28^{th} hydrogen was changed into an ethyl group, and for the some step for the conversion

of $complex_{myr}$ into $complex_{palm}$, high fluctuation of <dU/dλ> over time after was observed in the 10 ns length TI simulations. The change of 28^{th} hydrogen into an ethyl group causes higher conformational changes than the removal and addition of charge. Thus, for the TI simulation for the conversion of myristate into palmitate, production step 2 requires a longer simulation time (Durrant 2012). The fluctuation of <dU/dλ> over time in Durrant's project was observed as greater in all three TI simulation steps. Thus, stabilization of energy changes within the system proceeds faster when a smaller pressure relaxation time is used.

4.2 THE DIFFERENCE BETWEEN THE BINDING FREE ENERGY OF MYRISTATE AND PALMITATE

To calculate ΔΔG, graphs of <dU/dλ> with respect to λ were prepared. For the TI simulations, where myristate was converted into palmitate, curves for the change in <dU/dλ> with respect to λ are smooth. A relatively smooth change in <dU/dλ> with respect to λ was observed in the case of the conversion of $complex_{myr}$ into $complex_{palm}$, where a charge was added to the ethyl group, the step 3. While the change in <dU/dλ> with respect to λ for the removal of the charge from the 28^{th} hydrogen of myristate in the conversion of $complex_{myr}$ into $complex_{palm}$ is smooth excluding λ values of 0.1 and 0.7. The reason for λ values of 0.1 and 0.7 not following the rest of the curve could be that there was not appropriate convergence to a constant value of the <dU/dλ> over time, or it may indicate that more λ points are needed around these values to map the curve properly. For the conversion of $complex_{myr}$ into $complex_{palm}$, when the 28^{th} hydrogen was changed into an ethyl group the change in <dU/dλ> with respect to λ does not form a smooth curve, therefore, it is likely to give a bigger error in the ΔΔG calculation.

In our study smooth curves of <dU/dλ> with respect to λ were observed for the conversion of myristate into palmitate for all three production steps. It is intriguing that TI simulations for the conversion of myristate into palmitate for production step 2 do not have the expected convergence; nevertheless the change of <dU/dλ> with respect to λ is represented as smooth curves. In the Durrant's replicate study the change of <dU/dλ> with respect to λ for the conversion of $complex_{myr}$ into $complex_{palm}$ does not form smooth curves (Durrant, 2012). Moreover, the change of <dU/dλ> with respect to λ for the conversion of $complex_{myr}$ into $complex_{palm}$ in the replicate study is more chaotic than was observed in our

study. Therefore, in this study the calculated ΔΔG should be more close to the ΔΔG value that could be obtained if the TI simulations could be run for longer.

Both my study and Durrant's study graphs of <dU/dλ> over time and graphs of <dU/dλ> with respect to λ show that the TI simulations for the conversion of $complex_{myr}$ into $complex_{palm}$ were probably not run for long enough. Therefore, two ΔΔG values with the aim to predict ΔΔG value, which could be obtained when TI simulations run long enough, were calculated. The calculated ΔΔG values using the first method (described on page 61) after 10 ns and 14 ns TI simulations are -9.090 kJ/mol and -8.816 kJ/mol, respectively. Using the second method (described on page 61) ΔΔG values are -7.573 kJ/mol and -8.795 kJ/mol. The difference between both methods after 10 ns and 14 ns is about 18.1% and 0.2%, respectively. The second method for ΔΔG calculation uses fewer integration steps, thus making it more precise. Moreover, in the $complex_{myr}$ system greater energy changes were observed after production step 2 for 10 ns TI simulations. Thus, the first method can have higher calculation error and differs more.

Our results show that ΔΔG is about -9 kJ/mol. Therefore, even whilst the 14 ns long TI simulations of $complex_{myr}$ conversion into $complex_{palm}$ do not indicate appropriate convergence; the calculated ΔΔG values vary little as the simulation is extended.

In the replicate study, ΔΔG calculated using the second method was -9.164 kJ/mol after 10 ns TI simulation. As was mentioned above ΔΔG calculated by Durrant (2012) will a have higher uncertainty than our calculated ΔΔG due to higher energy fluctuation.

Our calculated ΔΔG shows that $\Delta G_M > \Delta G_P$. Thus, palmitate binds stronger to I-FABP than myristate does. This was reported by Velkov et al. (2005 and 2007). Therefore, our calculated ΔΔG shows that our system is qualitatively comparable with experimental data. Nevertheless, ΔΔG is about five to seven times smaller than the ΔΔG calculated from the

experimental K_d values. These K_d values were obtained in experiments under varied conditions (Velkov et al., 2005, Velkov et al., 2007), resulting in varied experimental ΔΔG values. Thus, it could be assumed that the difference between our ΔΔG value and experimental ΔΔG values is due to different conditions. However, the difference between experimental ΔΔG values is small, while the difference between our calculated and experimental ΔΔG is about 5-7 times higher. Thus, it is likely that this difference is due to insufficient sampling or inadequate force fields within our simulated system. Therefore, our system is not quantitatively adequate and requires optimization. In further research an improvement in the system is required so that ΔΔG is close to experimental values.

When the simulated system is optimized, TI simulations for artificially restricted myristate can be run. . The optimized system and TI integration data for genuinely flexible and artificially restricted myristate can be further used to calculate ΔΔG. The difference between the values can give an idea as to whether flexibility of FAs is important for stronger binding to FABP proteins.

4.3 INVESTIGATING THE FLEXIBILITY OF FREE MYRISTATE IN SOLUTION AND OF MYRISTATE BOUND TO I-FABP

As our system is qualitatively comparable to the experimental data, MD simulations were performed, and conformations of myristate were grouped into clusters. Clustering results allowed suggestions to be made about potential bond and angle restrictions that can be used in further investigation.

4.3.1 Clustering by the average link algorithm

Grouping conformations of myristate by RMSD (set by default as < 0.6 Å) into distinct sets from one two five it was observed that myristate free in solution has one dominant conformation, while bound myristate forms two large clusters. Grouping myristate conformations into five clusters the biggest cluster of free myristate consists of 65.2% while two biggest clusters of bound myristate consist of 45.8% and 32.5% from the total count of myristate conformations (Figure 3.6). It could be true that the formation of a dominant cluster while myristate is free in solution is essential for myristate binding to I-FABP. Moreover, the dominant conformation of free myristate was found as highly similar to the myristate conformation when it binds to I-FABP (Figure 3.8). Therefore, the formation of a dominant conformation might be a driving force for myristate's strong binding to I-FABP.

Although, higher similarity was observed comparing five bound myristate clusters between them (described on pages 66 and 67), the sum of conformation of two biggest clusters for free myristate is bigger than it is for bound myristate, respectively, 945 and 783 from total 1000 myristate conformations. From this comparison it is hard to justify whether conformations of bound myristate are more similar than conformations of free myristate. Moreover, the higher similarity could suggest, but cannot indicate whether free or bound myristate is more rigid, thus clustering by dihedral angles was performed.

4.3.2 Clustering by dihedral angles

Clustering by dihedral angles was performed for each dihedral of free myristate in solution and bound myristate, using 36 bins. The angle distribution for each dihedral is represented in graphs (Figure 3.9 – 3.10). This dihedral analysis confirms that free myristate differs in angle range from bound myristate only at the 1st, 2nd and 3rd dihedrals (Figure 3.9 – 3.10). The 1st, 2nd and 3rd dihedrals are positioned at the head group of myristate. It was reported that FAs binds to FABP protein with their head groups while their aliphatic tail group moves freely (Woolf and Tychko, 1998). The same was observed in our clustering results.

The preferred angle range at the 1st, 2nd and 3rd dihedrals of myristate becomes more concise upon myristate binding to I-FABP indicating that free myristate is more flexible at these dihedrals than bound myristate. Important to note that myristate binding to I-FABP as well leads to a change in the angle distribution at the 2nd dihedral angle.

In the replicate study, MD simulations using the same simulation conditions were performed for palmitate. Clustering by dihedral angles was performed for each dihedral of free palmitate in solution and bound palmitate, using 36 bins (Durrant, 2012). The structure of palmitate differs from myristate by an additional ethyl group instead of the myristate's 28th hydrogen.

Comparing the angle distribution for free myristate and free palmitate (Durrant, 2012) no high differences in the preferred angle ranges were observed. However, comparing the angle distribution for bound myristate and bound palmitate different preferred angle ranges at the 2nd, 4th to 6th dihedrals were observed. At the 4th and 6th dihedrals bound palmitate has only one preferred angle range while bound myristate has three preferred angle ranges. The dominantly preferred angle range for bound myristate and the only angle range for bound

palmitate are centred at 175º. The additional angle ranges for bound myristate indicate that at the 4th and 6th dihedrals bound myristate is more flexible than bound palmitate.

At the 2nd and 5th dihedrals the preferred angle ranges are shifted differently for bound myristate and bound palmitate. At the 2nd dihedral bound palmitate has only one dominate preferred angle range centred at -70º, while bound myristate forms two, the dominantly preferred angle range is centred at 55º and the second preferred angle range at -70º. At the 5th dihedral, bound palmitate has three preferred angle ranges, the dominant at -75º, and the other two at 50º and 180º. While bound myristate forms two preferred angle ranges, the dominant at 175º and the second at -70º.

Crystal structures of FA interaction with FABPs show that the crystal structure of the protein-ligand complex indicated that FA may penetrate FABP with either aliphatic tail (Stewart, 2000, Zimmerman and Veerkamp, 2002, Thompson et al., 1997) or carboxylate (Stewart, 2000, Zimmerman and Veerkamp, 2002, Wolfrum and Spener, 2000, Haunerland and Spener, 2004, Sacchettini et al., 1989). In the most cases when FA penetrates FABP either with aliphatic tail or carboxylate FA are half buried in the FABP, but it can be buried in the FABP fully. For example, Tsfadia et al. (2007) showed that palmitate penetrates Toad-liver FABP with the carboxylate group. At the first palmitate is half buried into Toad-liver FABP, the carboxylate group is interacts only with Q56. Further, the carboxylate penetrates deeper into the Toad-liver FABP cavity there it becomes stabilized by Q56 and H58 (Tsfadia et al., 2007).

Palmitate is longer than myristate by an additional ethyl group. As was mentioned above there is a difference in preferred dihedral angle ranges for myristate and palmitate. Thus, an additional ethyl group in palmitate should play a role in this difference. It might be that due to the additional group the tail of palmitate interacting with the I-FABP and is forced

to penetrate I-FABP deeper than myristate does. Although this has not been investigated, higher restriction and difference in the angle distribution comparing to myristate might be explained by the deeper palmitate penetration into I-FABP. Thus, it could be, due to this palmitate binds stronger to I-FABP than myristate does. However, the extra carbon atoms should increase the van der Waals interaction and the hydrophobic effect to enhance binding, which should improve the binding affinity. Thus, it might be that the bigger flexibility of free palmitate in comparison to free myristate does not play a role in the stronger binding to I-FABP.

4.3.2.1 The change in myristate flexibility during its binding to I-FABP

H/H_{max} was used as a measure of dihedral flexibility. Dihedrals of free myristate were arranged in order with increasing flexibility: $5^{th} \approx 9^{th} < 6^{th} = 10^{th} \approx 7^{th} = 3^{rd} \approx 8^{th} \approx 11^{th} \approx 4^{th} = 12^{th} \approx 2^{nd} < 1^{st}$, while dihedrals of bound myristate were arranged in order with increasing flexibility: $10^{th} < 8^{th} \approx 2^{nd} = 5^{th} = 12^{th} \approx 1^{st} = 9^{th} \approx 3^{rd} < 11^{th} < 7^{th} < 4^{th} < 6^{th}$. H/H_{max} for free myristate ranges from 0.60 to 0.88 at different dihedrals. Similar H/H_{max} values from 0.58 to 0.89 were calculated for free palmitate in the replicate study (Figure 3.11 (a)); indicating that flexibility of myristate at each dihedral is different. The most flexible free myristate and free palmitate are at the 1^{st} dihedral. This was expected, as the 1^{st} dihedral angle is formed between O-C_1 and C_2-C_3, where C_1 is bound to two oxygens which rotate freely around C_1-C_2 bond.

Myristate and palmitate binds to I-FABP with the head group. Thus, it was expected the head group will be the most rigid, when FAs binds to I-FABP. However, bound myristate is the most rigid at the 10^{th} dihedral, where H/H_{max} is 0.54. H/H_{max} for the 1^{st} and 2^{nd} and 3^{rd} dihedrals is just a bit lower than the average $H/H_{max} = 0.62$. In comparison bound palmitate is

the most rigid at the first four dihedrals (Figure 3.11 (a)), indicating that palmitate movements are restricted at these dihedrals by the interaction with the protein.

Although, the bound myristate is the most rigid at the 10^{th} dihedral, the decrease in flexibility when myristate binds to I-FABP is at the first dihedral. Dihedrals of myristate were arranged of the decreasing change of flexibility: $11^{th} \approx 3^{rd} < 10^{th} = 8^{th} \approx 12^{th} \approx 2^{nd} < 1^{st}$. It is intriguing that when myristate binds to I-FABP it becomes more flexible at the 6^{th} dihedral. At the 4^{th}, 5^{th}, 7^{th} and 9^{th} dihedrals no change in flexibility was observed. In the replicate study palmitate had the largest decrease in flexibility at the 1^{st}, 2^{nd} and 4^{th} dihedrals, the largest increase in flexibility at the 14^{th} and no change was observed at 5^{th}, 8^{th}, 10^{th}, 12^{th} and 13^{th}.

4.3.2.2 Correlation coefficient

Correlation coefficients for each dihedral angle were calculated to compare the similarity of the angle distribution between free and bound myristate. Dihedrals were arranged in order of increasing similarity between their free and bound myristate angle distribution is: $2^{nd} < 1^{st} < 6^{th} < 4^{th} < 3^{rd} \approx 7^{th} = 12^{th} \approx 5^{th} = 8^{th} \approx 9^{th} = 11^{th} \approx 10^{th}$. The highest difference between the free and bound myristate angle distribution was observed at the 1^{st} and 2^{nd} dihedrals, with correlation coefficients of 0.24 and 0.11, respectively, reflecting that myristate changes its preferred angle range for these dihedrals when it binds to I-FABP. At the other dihedrals, high similarity between free and bound myristate angle distribution was observed. The correlation coefficient at these dihedrals is > 0.94; except for the 4^{th} and 6^{th} it is 0.91 and 0.88, respectively.

During binding to I-FABP myristate adapts to the most energetically favorable conformation. An important fact is that the conformation of free myristate is very similar to the conformation when it is bound to I-FABP. The biggest differences between the free and bound myristate angle distribution and the biggest increase in the flexibility was observed at

the 6th dihedral, excepting the 1st and 2nd dihedrals which have already been discussed. The correlation coefficients calculated for palmitate dihedrals in the replicate study were reported as 0 for the 2nd and 5th and as 0.55 for the 1st indicating that there were no similarities between free and bound palmitate angle distribution at these dihedrals. For the other dihedrals the correlation coefficients were < 0.85, except the 10th and 13th dihedrals, where it was 0.77 and 0.81, respectively (Figure 3.11 (b)).

Comparing the correlation coefficients for myristate and palmitate, myristate has half smaller correlation coefficient at the 1st dihedral, indicating that the 1st dihedral of myristate undergoes higher conformational changes than it does in palmitate. The correlation coefficient for myristate and palmitate at the 2nd dihedral is 0.11 and 0.01, respectively (Figure 3.11 (b)). However, at the 2nd dihedral palmitate has a smaller correlation coefficient, a shift in the preferred angle range been observed for both FAs upon their binding to I-FABP. The correlation coefficient at the 5th dihedral for myristate indicates that the conformations of bound and free myristate are very similar at this dihedral. While the correlation coefficient at the 5th dihedral for palmitate indicates that they have no similarities. It might be that the changes at the 5th dihedral of palmitate are required to maintain an energetically favorable conformation for binding to I-FABP. At the 10th dihedral, the correlation coefficients for myristate and palmitate are 0.99 and 0.77, indicating that palmitate undergoes bigger conformational changes at this dihedral. However, the changes between conformation of free and bound palmitate are small. The correlation coefficients at the other dihedrals of myristate and palmitate are similar.

4.3.2.3 Is flexible FA stronger binding to I-FABP putative?

Comparing our myristate dihedral clustering and palmitate clustering by Durrant (2012) was suggested that due to two extra carbon atoms palmitate penetrates might penetrate I-

FABP deeper resulting in the higher torsion restriction and difference in the angle distribution Palmitate becomes more rigid at the first four dihedrals, while myristate only at the first two. It might be that palmitate restriction at these four dihedrals causes the increase in the rigidity in the remaining molecule. As well it might be that the dihedral changes are due to the different binding conformations of myristate and palmitate.

Myristate does not penetrate I-FABP as deep. Thus, the dynamics of the torsions in the head group of the bound myristate and bound palmitate might be different causing the palmitate to bind stronger to I-FABP. . As well it might be that palmitate binds stronger to I-FABP due to additional ethyl group allows palmitate to penetrate I-FABP deeper and not due to an increase in rotatable bonds. Thus, it might indicate that flexible FA strong binding to I-FABP is putative. It is assumed that FAs may have the best fit in the I-FABP cavity, which allows them to bind stronger to the protein than rigid ligands do.

4.4 FURTHER STUDIES

Our clustering results by dihedral angle show that during myristate binding to I-FABP myristate became more rigid at the 1st and 2nd dihedrals and more flexible at the 6th. Therefore, the restriction of myristate should be considered at these dihedrals for the purpose to calculate the free energy effect of restriction. In the further studies TI simulations for conversion of artificially rigid myristate into flexible myristate should be performed and the difference of the free binding energy between artificially rigid and flexible myristate should be calculated ($\Delta\Delta G_{R-F}$).

The restriction of free myristate at the 1st dihedral to imitate the dominant conformation of bound myristate is expected to affect the loss of entropy. Artificially rigid myristate may lose less entropy as common dogma states and thus bind better. Therefore, a decrease in $\Delta\Delta G_{R-F}$ would show that introduced rigidity improves myristate binding to I-FABP. While, a positive $\Delta\Delta G_{R-F}$ value would indicate that genuinely flexible myristate binds to I-FABP stronger than artificially rigid myristate does.

As it was discussed before when myristate binds to I-FABP at most of the dihedrals it becomes rigid. However, the bound myristate is more flexible at the 6th dihedral than when it is free in solution. Thus, a restriction of bound myristate at the 6th dihedral to imitate the dominant conformation of free myristate could show if an increase in the flexibility at this dihedral supports bound myristate formed conformation similar to the conformation when it is free in solution. The importance of flexibility would be confirmed by a decrease in $\Delta\Delta G_{R-F}$ value.

The 2nd dihedral as well should be considered for a restriction. When myristate binds to I-FABP, at this dihedral a shift in the preferred angle range was observed and myristate became more rigid around this torsion. . The restriction of bound myristate at this dihedral to

imitate the dominant conformation of free myristate by making the 2nd dihedral angle of free myristate 180° big could give us evidence of how the conformation of myristate affects its binding to I-FABP. It could be true that the dominant conformation of free myristate is important for its interactions with the protein. Nevertheless, the restriction may not lead to binding. The binding to I-FABP of the artificially restricted myristate can be simulated even if the restriction would not lead to the myristate and I-FABP binding in an experimental system. Therefore, the change in $\Delta\Delta G_{R-F}$ would show how big the effect of the conformational restriction on the binding of myristate and I-FABP is.

4.5 SUMMARY

TI simulations of the conversion of myristate into palmitate in solution indicate a rapid convergence of <dU/dλ>. TI simulations of the conversion of $complex_{myr}$ into $complex_{palm}$ proceeded with greater conformational changes within the system. Therefore, these TI simulations need to continue until the energy of the system becomes constant for all λ. In comparison to the Durrant's study, where pressure relaxation time was 1 ns, pressure relaxation time 0.2 ns used in our studies allows for faster convergence of TI simulations.

The calculated ΔΔG is around -9 kJ/mol for TI simulations run for 10 ns and 14 ns. Although the curves of <dU/dλ> for individual λ do not appear to completely converge, the calculated ΔΔG values vary little as the simulation is extended. Subsequently minimal ΔΔG changes are expected after further TI simulations are continuation.

Our ΔΔG is quantitatively different from the experimental data (Velkov et.al. 2005, Velkov et.al. 2007). This could be due to incomplete sampling (particularly of the slower protein motions), due to inadequacies in the force field or indeed inadequacies in the conditions used, although the first two reasons are more likely. Nevertheless, the negative value of ΔΔG, calculated from both ours and the Velkov et al. (2005 and 2007) data shows that palmitate binds to I-FABP stronger than myristate does. Therefore, the system can be used as a qualitative model in the investigation of the role of ligand flexibility in the binding of myristate to I-FABP. Nevertheless, optimization of the system would be advisable to make it match the experimental data more closely.

Clustering by the average link algorithm showed that both free myristate and bound myristate have a dominant conformation. Whilst binding to I-FABP, myristate becomes more rigid close to the head group. Change in conformation was also observed in the head group accompanied by a change in conformational preference.

At the 6th dihedral bound myristate is the most flexible, but at the 10th dihedral it is the most rigid. However, the 1st and 2nd dihedrals experience the biggest decrease in flexibility, upon binding I-FABP compared to the others dihedrals.

It is assumed that the more flexible palmitate may bind stronger to I-FABP due to its ability to penetrate I-FABP more than myristate can. Thus, it might indicate that flexible FA strong binding to I-FABP is putative. Thus, it might be that the structure of I-FABP and the structure of the protein cavity are favorable for FAs penetration and strong binding to I-FABP.

The restriction of the rotation around a dihedral *in silico* will lead to a measurable change in $\Delta\Delta G_{R\text{-}F}$. If we work on the assumption that flexibility could somehow be beneficial in the thermodynamics of myristate binding to I-FABP, then restricting myristate at the 1st dihedral should make $\Delta\Delta G_{R\text{-}F}$ more negative, but at the 6th dihedral should make $\Delta\Delta G_{R\text{-}F}$ less negative.

Thus, the simulations presented here provide a good basis for system optimization to qualitatively and quantitatively understand the role of flexibility in myristate binding to I-FABP, thereby resolving the potential anomaly that was described in the introduction.

5 REFERENCES

ABAD-ZAPATERO, C. (2007) Ligand efficiency indices for effective drug discovery. *Expert Opin Drug Discovery,* 2**,** 469-488.

ALDER, B. J. & WAINWRIGHT, T. E. (1959) Studies in Molecular Dynamics .1. General Method. *J Chem Phys,* 31**,** 459-466.

BENDER, H. (2003) *Molecular motions.* Amsterdam, Clackamas Community College, Available at: http://dl.clackamas.edu/ch106-01/molecula.htm (Accessed 18th September 2013).

BOEHR, D. D., DYSON, H. J. & WRIGHT, P. E. (2006) An NMR perspective on enzyme dynamics. *Chem Rev,* 106**,** 3055-3079.

BROOKS, B. R., BRUCCOLERI, R. E., OLAFSON, B. D., STATES, D. J., SWAMINATHAN, S. & KARPLUS, M. (1983) Charmm - a Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J Comput Chem,* 4**,** 187-217.

CAI, Y. & SCHIFFER, C. A. Decomposing the energetic impact of drug resistant mutations in HIV-1 protease on binding DRV. *J Chem Theory Comput,* 6**,** 1358-1368.

CASE, D. A., DARDEN, T., CHEATHAM, T. E., SIMMERLING, C., WANG, J., DUKE, R. E., LUO, R., WALKER, R. C., ZHANG, W., MERZ, K. M., ROBERTS, B. P., HAYIK, S., ROITBERG, A., SEABRA, G., KOLOSSVARY, I., WONG, K. F., PAESANI, F., VANICEK, J., LIU, J., WU, X., BROZELL, S. R., STEINBRECHER, T., GOHLKE, H., CAI, Q., YE, X., WANG, J., HSIEH, M. J., HORNAK, V., CUI, G., ROE, D. R., MATHEWS, D. H., SEETIN, M. G., C., S., BABIN, V., LUCHKO, T., GUSAROV, S., KOVALENKO, A. & KOLLMAN, P. A. (2010) Amber 11 Users' Manual. 306.

CHANG, C. E., CHEN, W. & GILSON, M. K. (2007) Ligand configurational entropy and protein binding. *Proc Natl Acad Sci U S A,* 104**,** 1534-1539.

CHAPLIN, M. (2010) *Water structure and science.* London, London South Bank University, Available at: http://frienergi.alternativkanalen.com/Water_Atoms.htm (Accessed 18th September 2013).

CHUANG, S., VELKOV, T., HORNE, J., PORTER, C. J. & SCANLON, M. J. (2008) Characterization of the drug binding specificity of rat liver fatty acid binding protein. *J Med Chem,* 51**,** 3755-3764.

CORNELL, W. D., CIEPLAK, P., BAYLY, C. I., GOULD, I. R., MERZ, K. M., FERGUSON, D. M., SPELLMEYER, D. C., FOX, T., CALDWELL, J. W. & KOLLMAN, P. A. (1995) A 2nd Generation Force-Field for the Simulation of Proteins, Nucleic-Acids, and Organic-Molecules. *J Am Chem Soc,* 117**,** 5179-5197.

DAVIS, J. H. & AGARD, D. A. (1998) Relationship between enzyme specificity and the backbone dynamics of free and inhibited alpha-lytic protease. *Biochemistry,* 37**,** 7696-7707.

DURANT, R. (2012) Importance of flexibility in fatty acid binding to I-FABP: An exception to conventional theory? *School of Biosciences.* Birmingham, University of Birmingham.

FRIEDMAN, R., NACHLIEL, E. & GUTMAN, M. (2006) Fatty acid binding proteins: Same structure but different binding mechanisms? Molecular dynamics simulations of intestinal fatty acid binding protein. *Biophys J,* 90**,** 1535-1545.

GENONI, A., MORRA, G., MERZ, K. M. & COLOMBO, G. (2010) Computational Study of the Resistance Shown by the Subtype B/HIV-1 Protease to Currently Known Inhibitors. *Biochemistry,* 49**,** 4283-4295.

HAUNERLAND, N.H. & SPENER, F. (2000) Fatty acid-binding proteins - insights from genetic manipulations. *Cell Mol Life Sci*, 59, 1096-1116.

HAYES J. M. & ARCHONTIS G. (2012) MM-GB(PB)SA Calculations of Protein-Ligand Binding Free Energies. *InTech*, ISBN: 978-953-51-0444-5.

HOUSECROFT, C.E. & CONSTABLE, E.C. (2006) *Chemistry.* Third ed. London, Pearson Ltd.

HORNAK, V., ABEL, R., OKUR, A., STROCKBINE, B., ROITBERG, A. & SIMMERLING, C. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins*, 63, 712-725.HENZLER-WILDMAN, K. A., LEI, M., THAI, V., KERNS, S. J., KARPLUS, M. & KERN, D. (2007) A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature,* 450**,** 913-927.

JORGENSEN, W.L., CHANDRASEKHAR, J., MADURA J.D., IMPREY, R.W. & KLEIN, M.L. (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys,* 79**,** 926-935.

KUNTZ, I. D., CHEN, K., SHARP, K. A. & KOLLMAN, P. A. (1999) The maximal affinity of ligands. *Proc Natl Acad Sci U S A,* 96**,** 9997-10002.

LASSILA, J. K. (2010) Conformational diversity and computational enzyme design. *Curr Opin Chem Biol,* 14**,** 676-682.

LEACH, A. R. (2001) *Molecular modelling: Principles and applications,* Dorset, Henry Ling Ltd.

LEVIN, L. B., GANOTH, A., AMRAM, S., NACHLIEL, E., GUTMAN, M. & TSFADIA, Y. (2010) Insight into the interaction sites between fatty acid binding proteins and their ligands. *J Mol Model,* 16**,** 929-938.

LEVIN, L. B. A., NACHLIEL, E., GUTMAN, M. & TSFADIA, Y. (2009) Molecular dynamics study of the interaction between fatty acid binding proteins with palmitate mini-micelles. *Mol Cell Biochem,* 326**,** 29-33.

LUCKE, C., ZHANG, F. L., RUTERJANS, H., HAMILTON, J. A. & SACCHETTINI, J. C. (1996) Flexibility is a likely determinant of binding specificity in the case of ileal lipid binding protein. *Structure,* 4**,** 785-800.

MANN, A. (2008) Conformational Restriction and/or Steric Hidrance in Medicinal Chemistry. IN WERMUTH, C. G. (Ed.) *The practice of medicinal chemistry.* Third ed. London, Elsevier Ltd.

MICHEL, J., FOLOPPE, N. & ESSEX, J. W. (2010) Rigorous Free Energy Calculations in Structure-Based Drug Design. *Mol Inf,* 29**,** 570-578.

MURRAY, C. W. & VERDONK, M. L. (2002) The consequences of translational and rotational entropy lost by small molecules on binding to proteins. *J Comput Aided Mol Des,* 16**,** 741-753.

NELSON, D. L. & COX, M. M. (2000) *Lehninger principles of biochemictry*. Third ed., New York, Worth Publishers.

PAGE, M.I. & JENCKS, W.P. (1971) Entropic contributions to rate accelerations in enzymic and intramolecular reactions and the Chelate effect. *Proc Natl Acad Sci U S A*,68, 1678-1683.

PONDER, J. W. & CASE, D. A. (2003) Force fields for protein simulations. *Adv Protein Chem,* 66**,** 27-85.

REN, J. S., NICHOLS, C., BIRD, L. E., FUJIWARA, T., SUGIMOTO, H., STUART, D. I. & STAMMERS, D. K. (2000) Binding of the second generation non-nucleoside inhibitor S-1153 to HIV-1 reverse transcriptase involves extensive main chain hydrogen bonding. *J Biol Chem,* 275**,** 14316-14320.

SCOTT, W. R. P., HUNENBERGER, P. H., TIRONI, I. G., MARK, A. E., BILLETER, S. R., FENNEN, J., TORDA, A. E., HUBER, T., KRUGER, P. & VAN GUNSTEREN, W. F. (1999) The GROMOS biomolecular simulation program package. *J Phys Chem A,* 103**,** 3596-3607.

SEARLE, M. S. & WILLIAMS, D. H. (1992) The Cost of Conformational Order - Entropy Changes in Molecular Associations. *J Am Chem Soc,* 114**,** 10690-10697.

SHAO, J. Y., TANNER, S. W., THOMPSON, N. & CHEATHAM, T. E. (2007) Clustering molecular dynamics trajectories: 1. Characterizing the performance of different clustering algorithms. *J of Chem Theory Comput,* 3**,** 2312-2334.

SIMONSON, T., CARLSSON, J. & CASE, D. A. (2004) Proton binding to proteins: pK(a) calculations with explicit and implicit solvent models. *J Am Chem Soc,* 126**,** 4167-4180.

STEINBRECHER, T., MOBLEY, D. L. & CASE, D. A. (2007) Nonlinear scaling schemes for Lennard-Jones interactions in free energy calculations. *J Chem Phys,* 127**,** 214108.

TEAGUE, S. J. (2003) Implications of protein flexibility for drug discovery. *Nat Rev Drug Discovery,* 2**,** 527-541.

TSFADIA, Y., FRIEDMAN, R., KADMON, J., SELZER, A., NACHLIEL, E. & GUTMAN, M. (2007) Molecular dynamics simulations of palmitate entry into the hydrophobic pocket of the fatty acid binding protein. *Febs Lett,* 581**,** 1243-1247.

VELAZQUEZ-CAMPOY, A., LUQUE, I. & FREIRE, E. (2001) The application of thermodynamic methods in drug design. *Thermochim Acta,* 380, 217-227.

VELKOV, T., CHUANG, S., WIELENS, J., SAKELLARIS, H., CHARMAN, W. N., PORTER, C. J. & SCANLON, M. J. (2005) The interaction of lipophilic drugs with intestinal fatty acid-binding protein. *J Biol Chem*, 280, 17769-17776.

VELKOV, T., HORNE, J., LAGUERRE, A., JONES, E., SCANLON, M. J. & PORTER, C. J. (2007) Examination of the role of intestinal fatty acid-binding protein in drug absorption using a parallel artificial membrane permeability assay. *Chem Biol*, 14, 453-465.

VRIEND, G. (1990) What if: A molecular modelling and drug design program. *J Mol Graphics,* 8, 52-56.

WOOLF, T. B. & TYCHKO, M. (1998) Simulations of fatty acid-binding proteins. II. Sites for discrimination of monounsaturated ligands. *Biophys J*, 74, 694-707.

YU, L. P., ZHU, C. X., TSEDINH, Y. C. & FESIK, S. W. (1996) Backbone dynamics of the C-terminal domain of Escherichia coli topoisomerase I in the absence and presence of single-stranded DNA. *Biochemistry*, 35, 9661-9666.

Project 2:

STUDYING THE EFFECT OF THE BACTERIAL NDK PROTEINS ON HAEMOPOIETIC STEM CELLS

ABSTRACT

Leukaemia patients lack immunity, thus they are more responsive to bacterial infections. Moreover, infections are likely to worsen the leukaemia prognosis. It might be that bacterial invasion can affect leukaemia cells, by rendering the latter more aggressive, or even by stimulating cell proliferation.

Previous research has indicated that overexpression of NM23-H1 is related with poor prognosis of leukaemia patients. Moreover, additional rNM23-H1 is able to bind to more mature cells and indirectly promote leukaemia and healthy stem cell survival and proliferation. High structural similarity in NDK proteins may lead to functional similarity. Thus, additional bacterial NDK protein may indirectly promote leukaemia, healthy stem cell survival and proliferation as it was shown using rNM23-H1.

In our project we optimized purification of $CD34^{+ve}$ reaching 75.8 - 85.7% high purity. The positive control (rNM23-H1) failed to work; hence no suggestions can be made about rNDK effect on cord blood stem cells. The fail of positive control can be explained by possible rNM23-H1 activity loss due to protein aggregation after purification. Furthermore, the protocol of cord blood storage used for CB samples from previous studies was altered in our study. This change could affect $CD34^{-ve}$ and/or $CD34^{+ve}$ cell activity. Based on the latter, our study aimed to explore valuable directions for future research.

TABLE OF CONTENTS

TABLE OF FIGURES

ABBREVIATIONS

ALL	Acute lymphocytic leukaemia
AML	Acute myeloid leukaemia
BSA	Bovine serum albumin
CB	Umbilical cord blood
$CD34^{+ve}$/CM-EB	$CD34^{+ve}$ cells treated with conditioned media collected from treated $CD34^{-ve}$ cells with elution buffer
$CD34^{+ve}$/CM-EC	$CD34^{+ve}$ cells treated with conditioned media collected from treated $CD34^{-ve}$ cells with *E.coli* rNDK protein
$CD34^{+ve}$/CM-KP	$CD34^{+ve}$ cells treated with conditioned media collected from treated $CD34^{-ve}$ cells with *K.pneumoniae* rNDK protein
$CD34^{+ve}$/CM-NM	$CD34^{+ve}$ cells treated with conditioned media collected from treated $CD34^{-ve}$ cells with rNM23-H1
$CD34^{+ve}$/CM-SA	$CD34^{+ve}$ cells treated with conditioned media collected from treated $CD34^{-ve}$ cells with *S.aureus* rNDK protein
$CD34^{+ve}$/CM-SP	$CD34^{+ve}$ cells treated with conditioned media collected from treated $CD34^{-ve}$ cells with *S.pneumoniae* rNDK protein
CFU	Colony forming unit
CLL	Chronic lymphocytic leukaemia
CM	Conditioned media
CM_{dep}	Recombinant NDK protein depleted conditioned media
CM-EB	Conditioned media collected from $CD34^{-ve}$ cells treated with elution buffer
CM-EC	conditioned media collected from $CD34^{-ve}$ cells treated with *E.coli* rNDK protein
CM-KP	Conditioned media collected from $CD34^{-ve}$ cells treated with *K.pneumoniae* rNDK protein
CML	Chronic myeloid leukaemia
CM-NM	Conditioned media collected from $CD34^{-ve}$ cells treated with rNM23-H1 protein
CM-SA	Conditioned media collected from $CD34^{-ve}$ cells treated with *S.aureus* rNDK protein

CM-SP	Conditioned media collected from $CD34^{-ve}$ cells treated with *S.pneumoniae* rNDK protein
EB	Elution buffer
EBV	Epstein-Barr virus
GEC	Gingival epithelial cell
GLB	Gel loading buffer
GM-CSF	Granulocyte and macrophage colony stimulating factor
HPV	Human papillomavirus
HSC	Haemopoietic stem cell
HSCT	Haemopoietic stem cell transplantation
I-factor	Inhibitor factor
IL	Interleukin
IPTG	Isopropyl β-D-1-thiogalactopyranoside
KSHV	Kaposi's Sarcoma associated herpesvirus
LSC	Leukemia stem cell
LT-HSC	Long term haemopoietic stem cell
MACS	Magnetic activated cell sorting
M-CSF	Monocyte colony stimulating factor
MPP	Multipotent progenitor
NDK	Nucleotide diphosphate kinase
NM23	Non metastatic protein 23
NM23-H1	Non metastatic protein 23 homologues 1
PMB	Polymyxin B sulphate salt
PVDF	Polyvinilidene difluoride
rNDK protein	Recombinant nucleotide diphosphate kinase protein
rNDK-EC	*E.coli* rNDK protein
rNDK-KP	*K.pneumoniae* rNDK protein
rNDK-SA	*S.aureus* rNDK protein
rNDK-SP	*S.pneumoniae* rNDK protein
rNM23-H1	Recombinant non metastatic protein 23 homologues 1
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis

ST-HSC	Short term haemopoietic stem cell
TF	Transcription factor

1 INTRODUCTION

The immune system is an important biological system which acts with the aim to protect the body against infections caused by bacteria, parasites, fungi and viruses. Blood cells such as macrophages, neutrophils, mast cells, eosinophils, basophils, and natural killer cells are involved in defense of the body (Janeway et al., 2005).

Formation of normal haemopoietic cells is disrupted in leukaemia patients; thus leukaemia patients lack immunity (Wang and Dick, 2005) and are more responsive to bacterial infections (Bucaneve et al., 2007). Moreover, infections are likely to worsen leukaemia prognosis. It might be that bacterial invasion can affect leukaemia cells making them more aggressive or somehow the bacteria may support leukaemia cell proliferation.

The nucleoside diphosphate kinase (NDK) protein family includes is highly conserved and includes bacterial and eukaryotic proteins such as human non-metastatic protein 23 (NM23) group. Most of the work on human NDK protein (NM23 protein) in leukaemia has been carried out in a recent past by Okabe-Kado and Willems, who focused mainly on expression of human non-metastatic protein 23 (NM23) by haemapoietic cells and investigated the effect of additional NM23 proteins on healthy haemapoietic and leukemia cells. Since then has been shown that overexpression of NM23 homologue 1 (NM23-H1) is related with patient poor prognosis (Steeg et al., 1988, Yokoyama et al., 1996, Willems et al., 1998, Niitsu et al., 2000, Okabe-Kado et al., 2002, Okabe-Kado et al., 2009). Additional NM23 proteins can inhibit differentiation of healthy and leukaemia cells (Okabe-Kado et al., 1988, Okabe-Kado et al., 1992, Okabe-Kado et al., 1995) or alter cytokine induced differentiation (Willems et al., 2002), and promote acute myeloid leukaemia (AML) cell survival (Okabe-Kado et al., 2009). Lilly et al (2011) showed the most important finding in NM23 protein investigation, where additional recombinant (rNM23-H1) protein can bind to

more mature $CD34^{lo}/CD11b^{+ve}$ cells and indirectly promote survival and proliferation of healthy and leukaemia stem cells ($CD34^{+ve}$). Because bacterial and eukaryotic NDK proteins are highly conserved, we hypothesize that bacterial NDK proteins will act on leukaemia cells as it was observed by Lilly et al (2011) using the NM23-H1 protein.

1.1 HAEMOPOIESIS

The process of blood cell formation from haemopoietic stem cells (HSCs) is called haemopoiesis. Differentiation of HSCs promotes formation of various blood cell lineages (Figure 1.1). Blood cells are divided into red and white cells, where red cells incorporate the erythroid lineage. White cells are divided into myeloid and lymphoid cells (Hoffbrand et al., 2008).

1.1.1 Haemopoietic stem cells (HSCs)

HSCs have multipotent stem cell capacity, as they are able to differentiate into haemopoietic progenitor cells, which can further differentiate and maturate into various haemopoietic cells (Figure 1.1). Moreover, HSCs are capable of self-renewal. Thus, HSCs produce transit amplifying cells and then split into differentiated lineages (Hoffman et al., 2009). Progenitor cells differentiate and forms colony forming units (CFUs) (Figure 1.1). Further differentiation lead to mature blood cells. Despite progenitor cell ability to differentiate, they and their successors do not have self-renewal potential. Therefore, all blood cells can only be derived by HSC differentiation (Hoffbrand et al., 2008).

In murine bone marrow, the ratio of HSCs to the total amount of blood cells is one in 10,000 to 100,000. HSCs can be distinguished into long-term and short-term HSCs (LT-HSCs and ST-HSCs, respectively). Most of the HSCs are LT-HSCs that are. When LT-HSCs leave the steady state they become ST-HSCs that have the potential to differentiate into mature blood cells. LT-HSCs have the highest self-renewal potential. Moreover, *in vivo* studies have shown that ST-HSCs have limited self-renewal potential as they disappeared after 4-6 weeks (Coulombel, 2004).

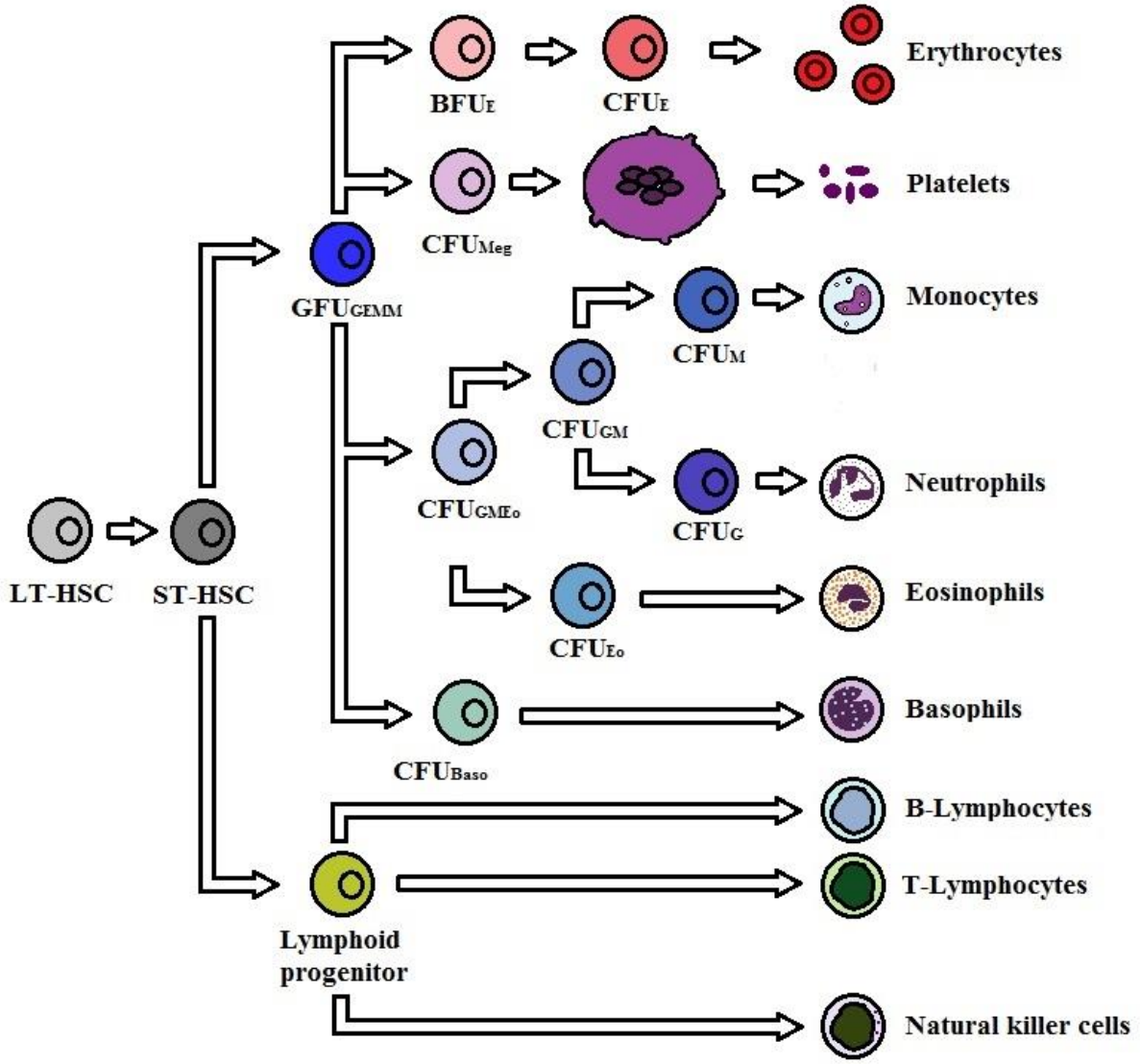


Figure 1.1 Formation of blood cell lines. LT-HSC and ST-HSC - long term and short term haemopoietic stem cells, BFU – burst forming unit, CFU – colony forming unit, Baso – basophil, E – erythroid, Eo – eosinophil, G – granulocyte, GEMM – granulocyte, erythroid, monocyte and megakaryocyte, GM – granulocyte, monocyte, GMEo – granulocyte, monocyte and eosinophil, M – monocyte, Meg – megakaryocyte. (Reproduced from Hoffman et al, 2009 and Hoffbrand et al, 2006)

The main site of haemopoiesis in adults is bone marrow. HSCs and progenitor cells differentiate here and usually do not leave the bone marrow. Nevertheless, they are able to exit the bone marrow as they are found in peripheral blood in low numbers. High migration of HSCs occurs in haematological malignancies (Hoffbrand et al., 2008, Hoffman et al., 2009).

Bone marrow is distinguished into two niches – endosteal and vascular (Figure 1.2). The endosteal niche is formed by osteoblasts, stromal fibroblasts and CXCL12-abundant reticular cells. These cells are involved in the regulation of HSC number and their function. HSCs found in endosteal niche are quiescent. Osteoblasts are maintaining HSC self-renewal. Migration of HSCs to a vascular niche leads HSCs to become active. Thus, in the vascular niche HSCs proliferate and differentiate (Buske et al., 2002, Wilson and Trumpp, 2006, Lilly, 2011).

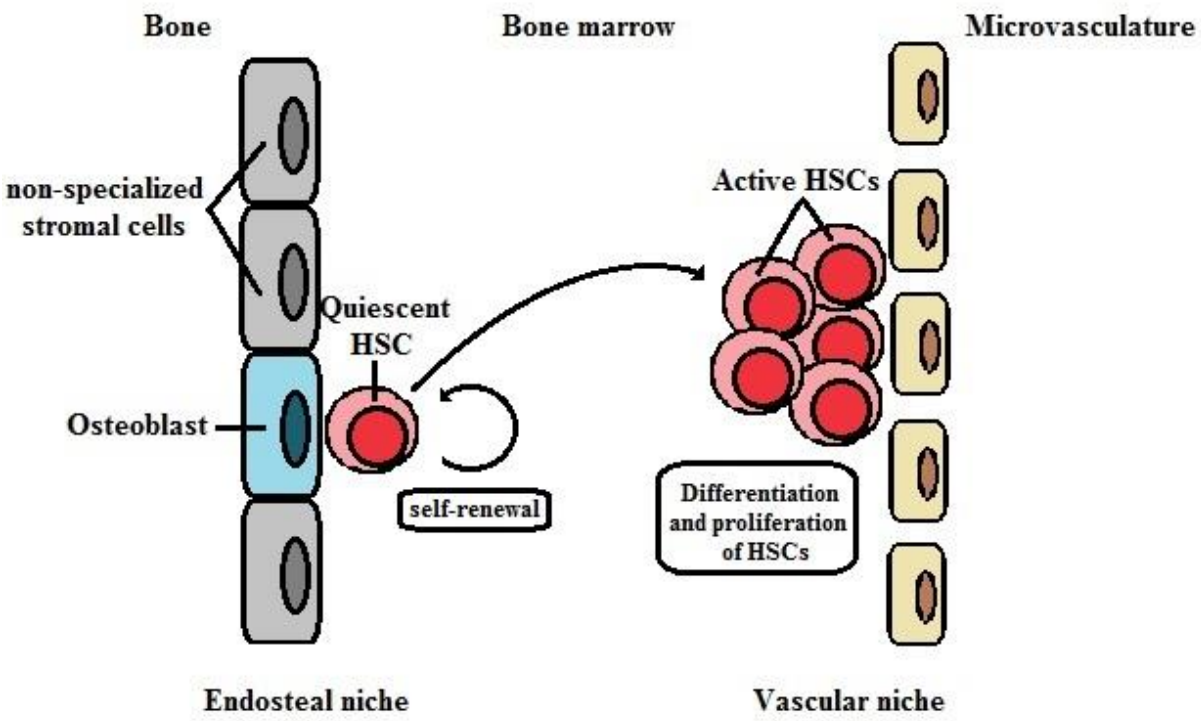


Figure 1.2 Bone marrow niches. (Reproduced from Wilson and Trumpp, 2006)

1.1.2 Leukaemia stem cells (LSCs)

Mutations within HSCs may lead HSCs to become cancer stem cells called leukaemia stem cells (LSCs). The important fact is that any cancer is a heterogenic cell population. Thus, within a cancer there are cancer stem cells and more mature cancer cells (Heppner and Miller, 1993, Dick, 2008). Cytotoxic chemotherapy is one cancer treatment where chemotherapeutic agents kill rapidly dividing cells. As quiescent LSCs do not divide, they may survive chemotherapy and restore depleted leukaemia cell heterogeneity (Dick, 2008, Sehl et al., 2009). Bacteria may invade leukaemia patients and can affect LSCs making them more aggressive and more resistant to chemotherapeutic agents, thus causing leukaemia to reoccur. However, bacteria can invade leukaemia patient after chemotherapy. Thus, bacteria may support LSC proliferation and relapse of leukaemia.

1.1.3 Haemopoietic cell molecular markers

Cell molecular markers are antigens of which the expression changes with cell differentiation (Dick, 2008). Therefore, molecular markers allow us to distinguish haemopoietic cells at the various development stages. Unfortunately, there is no unique marker for HSCs, progenitor cells or even mature haemopoietic cells. To distinguish these cells combination markers can be used (Hoffman et al., 2009).

Some of the cell molecular markers are expressed only in immature cells, whilst the others are found only in the more mature cells. For example, human HSCs and multipotent progenitor (MPP) cells express CD34. It is highly expressed in the most immature HSCs, its expression lowers with stem cell maturation (Shizuru et al., 2005). While CD38 marker is absent in an early stage of heamopoiesis and its expression is observed on common myeloid progenitor cells (CMPs) (Wang et al., 2012).

Cell molecular markers are species dependent. The markers that are commonly used to identify mouse and human HSCs are shown in the Figure 1.3 (a). LSCs, as well as HSCs, express CD34 marker and lacks CD38. However, there is a difference in expressed markers by LSCs and HSCs (Figure 1.3). LSCs lack CD90 and CD117 cell markers, which are expressed by HSCs. Moreover, CD123 known as well as interleukin-3 receptor α (IL-3Rα) is expressed only by LSCs (Wang and Dick, 2005). It might be that receptor IL-3Rα is required for LSCs abnormal behavior and life cycle maintenance.

Different types of mature haemopoietic cells may express the same markers such as CD11b as well as an individual marker for the cell lineage. CD11b is expressed by monocyte-macrophage lineage (Grigoriadis et al., 2010) as well by natural killer cells and T-cells (Moreau et al., 2009).

(a)

Mouse haemapoietic cell markers

HSC/LT-HSC	$CD34^{-ve}$ Lin^{-ve} $c\text{-}Kit^{+ve}$ $Sca1^{+ve}$ $CD150^{+ve}$
ST-HSC/MPP	$CD34^{+ve}$ Lin^{-ve} $c\text{-}Kit^{+ve}$ $Sca1^{lo}$ $CD150^{-ve}$

Human haemapoietic cell markers

HSC	$CD34^{+ve}$ $CD38^{-ve}$ Lin^{-ve} $CD90^{+ve}$ $CD45RA^{-ve}$ $CD49f^{+ve}$
MPP	$CD34^{+ve}$ $CD38^{-ve}$ Lin^{-ve} $CD90^{-ve}$ $CD45RA^{-ve}$ $CD49f^{-ve}$
CMP	$CD34^{+ve}$ $CD38^{+ve}$ Lin^{-ve} $CD45RA^{-ve}$ $CD123^{med}$ $CD135^{+ve}$
CLP	$CD34^{+ve}$ $CD7^{+ve}$ $CD10^{+ve}$

(b)

Haemapoietic and leukaemia stem cell markers

HSC	Lin^{-ve} $CD34^{+ve}$ $CD38^{-ve}$ $CD90^{+ve}$ $CD117^{+ve}$ $CD123^{-ve}$
LSC	Lin^{-ve} $CD34^{+ve}$ $CD38^{-ve}$ $CD90^{-ve}$ $CD117^{-ve}$ $CD123^{+ve}$

Figure 1.3 Expression of haemopoietic cell markers, where LT-HSC and ST-HSC - long term and short term haemopoietic stem cells, MPP - multipotent progenitors, LSC – leukaemia stem cells, CMP = common myeloid progenitor cells and CLP – common lymphoid progenitor cells (a) Mouse (adapted from Wognum and Szilvassy (2013)) and human haemopoietic cell markers (adapted from Wang et al, 2012 (b) comparison of cell markers between LSCs and HSCs (cell molecular markers are adapted from Wang and Dick, 2005)

1.2 LEUKAEMIA

Leukaemia is cancer of the blood cells, which manifests itself by an increase in functionally immature blood cells (Wang and Dick, 2005). Accumulation of the immature cells in the bone marrow causes a decrease in the healthy cell development. Thus, leukaemia patients are immunocompromised. Moreover, the excess of leukaemia cells may exit the bone marrow, and they spread to the other organs.

There are four main types of leukaemia: acute and chronic lymphocytic leukaemia (ALL and CLL, respectively), and acute and chronic myeloid leukaemia (AML and CML, respectively). Other rarer leukaemias also exist, e.g. hairy cell leukaemia, T-cell prolymphocytic, large granular lymphocytic, adult T-cell leukaemia, etc. (Hoffbrand et al., 2008).

1.2.1 Acute myeloid leukaemia (AML) and its treatments

AML is a cancer of myeloid blood lineages. It proceeds as abnormal white blood cell growth. Primary AML occurs in previously healthy patients whilst secondary AML develops from myelodysplastic syndromes (MDS) or other haemopoietic diseases. AML patients are assigned to the secondary AML group if they had any exposure to carcinogens, radiation or they have reoccurring AML (Rowe, 2002).

Common treatments for leukaemia including AML are chemotherapy, radiotherapy, blood transfusion, bone marrow transplants or other drug treatments. Treatment method depends on the type of AML as well on the age of a patient (Burnett et al., 1998))

1.2.2 Bacterial infections in the AML patients

Human beings are continuously exposed to harmful microbiological organisms. In a healthy person, the immune system controls the extent and duration of the infection. As the immune system of AML patients is depleted, they are at a high risk of bacterial, viral or fungal infections (Janeway et al., 2005). Moreover, infections are likely to worsen the AML prognosis. From clinical data in the period from 1980 until 2005, Buncaneve et al (2007) quantified that bacterial infections are responsible for 30-40% of all febrile episodes in acute leukaemia and haemopoietic stem cell transplantation (HSCT) patients. Moreover, 30% of all infections were caused by infection of the bloodstream (Bucaneve et al., 2007).

Bacterial infection can be caused by gram-negative and gram-positive bacteria. These were accounted as 10% and 15%, respectively, of total febrile episodes in acute leukaemia and HSCT patients, respectively. The most common gram-negative bacteria causing an infection in leukaemia patients are *Escherichia coli* (Kern et al., 2005, Gomez et al., 2003, Buncaneve et al., 2007), *Pseudomonas* spp such as *Pseudomonas aeruginosa* (Kang et al., 2003), *Enterobacter cloacae* (Tancrede and Andremont, 1985) and *Salmonella typhimurium* (Helms et al., 2002, Tancrede and Andremont, 1985). Most common gram-positive bacteria are *Staphylococcus* spp such as *Staphylococcus aureus* (Madani, 2000, Marty et al, 2006, Buncaneve et al., 2005) and *Streptococcus* spp such as *Streptococcus pneumonia* (Madani, 2000, Bucaneve et al., 2007). *E.coli* and *Staphylococcus* strains such as *Staphylococcus epidermidis* are not pathogenic to healthy humans, yet may cause infection in immune-suppressed patients.

Prophylaxis of bacterial infection in the AML patient treatment includes antibiotics, antiseptics for bathing, mouthwashes and a “clean diet”. Patients with neutropenia, on anti-cancer chemotherapy, and severely immunocompromised patients to prevent bacterial and

viral infections are usually kept in a reverse-barrier room. The amount of protection required varies with the patient (Hoffbrand et al., 2008, Schimpff et al., 1975).

Prophylaxis is very important, but it does not protect leukaemia patient from bacterial infection 100%, and it does not cure leukaemia. Nevertheless, it improves patient prognosis, for example, prophylaxis of acute leukaemia and HSCT patients using fluoroquinolone was associated with a reduction in mortality of 48% in causes, and a reduction in infection-related mortalities of 68% (Bucaneve et al., 2007).

In conclusion, bacterial infections are likely to occur in leukaemia patients as their immune system is suppressed. Bacterial infection prophylaxis using antibiotics has been shown to reduce mortality in leukaemia patients. Therefore, a number of questions arise; do bacterial infections affect or support the development of leukaemia? Is there a united mechanism for different bacterial strains to affect leukaemia?

1.3 NUCLEOSIDE DIPHOSPHATE KINASE (NDK) PROTEINS

NDK proteins are enzymes (phosphotransferases), which reversibly catalyze the exchange of terminal phosphate groups between nucleoside diphosphate and nucleoside triphosphate (Berg et al., 2002).

NDK proteins are highly conserved. Human NDK proteins are named as non-metastatic proteins 23 (NM23) because at their discovery; the time expression of NM23-H1 protein was found related to the suppression of metastasis (Steeg et al., 1988). Ten homologues of NM23 have been identified, and named as NM23-H1, NM23-H2, NM23-H3, NM23-H4, NM23-H5, NM23-H6, NM23-H7, NM23-H8, NM23-H9 and NM23-H10 (Bilitou et al., 2009). The sequence of NM23 homologues is shown in figures 1.4. The biggest sequence similarity (88.2%) is between NM23-H1 and NM-23-H2, while 65.8% and 55.3% of the NM23-H1 sequence is identical to the sequence of NM23-H3 and NM23-H4, respectively (Figure 1.4).

Not only human NDK homologues have high sequence identity. Eukaryotic and bacterial NDK proteins are also highly conserved. For example, *E.coli* NDK protein has 57% sequence identity with its homologue in *Myxococcus xanthus*, 45% with *Dictyostelium discoideum* (Hama et al., 1991) and 95% with *K.pneumoniae* NDK protein (Figure 1.5). NM23-H2 possess 45% and 67% sequence identity with *Mycobacterium tuberculosis* (Chen et al., 2002) and *Dictyostelium* NDK protein (Dumas et al., 1992), respectively. NM23-H1 has 43%, 44%, 47% and 53% sequence identity with *E.coli, K.pneumoniae, S.pneumoniae, S.aureus* NDK proteins, respectively (Figure 1.5).

Eukaryotic and bacterial NDK proteins are structurally very similar. The Killer of prune (Knp) loop, α helixes and β sheets are distinguished in the NDK protein tertiary structure. In figure 1.6, they are matched on to the *E.coli, K.pneumoniae, S.pneumoniae,*

S.aureus NDK protein and NM23-H1 sequences. The most conserved regions are the Knp loop, β1 and α5. The Kpn loop has an important role in kinase activity of the NDK protein (Bilitou et al., 2009). Thus, structural similarities in-between bacterial and human NDK proteins may also mean that bacterial and human NDK proteins have functional similarity. Thus, bacteria may affect leukaemia cells in a similar way.

H3

	10					15						20					25						30					35					40	
H1									M	A	N	C	E	R	T	-	F	I	A	I	K	P	D	G	V	Q	R	G	L	V	-	G	
H2									•	•	•	L	•	•	•	-	•	•	•	•	•	•	•	•	•	•	•	•	•	•	-	•	
H3	F	A	N	L	F	P	A	A	C	T	G	A	H	•	•	•	-	•	L	•	V	•	•	•	•	•	•	•	R	•	•	-	•	
H4		V	R	H	G	S	G	G	P	S	W	T	R	•	•	•	-	L	V	•	○	•	•	•	•	•	•	•	○	•	•	-	•	
H5		M	E	I	S	M	P	P	□	Q	I	Y	V	•	K	•	-	□	A	I	•	•	•	•	I	•	D	K	E	-	-	-	E	
H6			M	A	◊	I	L	R	S	P	Q	○	Δ	Q	L	•	-	□	◊	L	•	•	•	•	A	•	A	H	P	•	I	L	◊	
H7									•	N	H	S	•	•	F	V	•	•	•	E	W	Y	•	P	N	⌂	S	L	•	R	R	Y	
H8			⌂	⌂	◊	K	K	⌂	E	V	⌂	L	Q	T	V	I	-	N	N	Q	S	L	W	•	E	M	L	Q	N	K	-	-	•	
H9			⌂	○	◊	□	#	□	L	○	V	V	D	V	Q	G	-	W	C	G	P	C	K	P	V	•	S	F	Q	#	M	x	I	
H10	M	G	C	F	○	A	#	⌂	R	K	•	D	K	•	S	R	-	P	E	N	x	E	E	R	x	K	•	Y	S	W	D	Q	R	

	60					65						70					75						80					85					90	
H1	F	M	Q	A	S	E	D	L	L	K	E	H	Y	V	D	L	K	D	R	P	F	F	A	G	L	V	K	Y	M	H	S	G	P	
H2	•	L	R	•	•	•	E	H	•	•	Q	•	•	I	•	•	•	•	•	•	•	•	P	•	•	•	•	•	•	N	•	•	•	
H3	L	V	•	•	•	•	Δ	•	•	R	•	•	•	A	E	•	R	E	•	•	•	Y	G	R	•	•	•	•	•	A	•	•	•	
H4	M	Δ	•	•	P	•	S	V	•	A	•	•	•	Q	•	•	○	R	K	•	•	○	Δ	A	•	I	R	•	•	S	•	•	•	
H5	K	Δ	Δ	L	•	P	Δ	Q	C	S	N	F	•	•	○	K	Y	G	□	M	•	•	Δ	N	•	T	A	•	•	□	•	•	•	
H6	E	Δ	L	W	R	K	Δ	D	◊	Q	R	◊	•	R	○	H	E	◊	•	F	•	○	Q	○	•	•	E	F	•	○	•	•	•	
H7	D	○	K	N	H	R	T	F	•	•	⌂	T	K	Y	•	N	L	H	I	E	D	L	F	I	G	N	•	V	N	V	F	S	R	
H8	Q	P	⌂	F	⌂	⌂	L	K	N	E	L	N	E	D	○	I	x	x	F	A	V	A	E	□	D	x	I	x	T	L	Q	P	F	
H9	x	○	⌂	E	K	Y	R	G	K	C	•	P	T	F	L	F	◊	A	x	K	x	E	•	L	S	D	⌂	D	E	C	V	x	H	
H10	V	G	Δ	◊	□	G	x	□	A	G	Δ	Q	-	×	×	#	Q	•	C	x	N	C	N	x	Y	□	F	×	-	•	•	A	T	

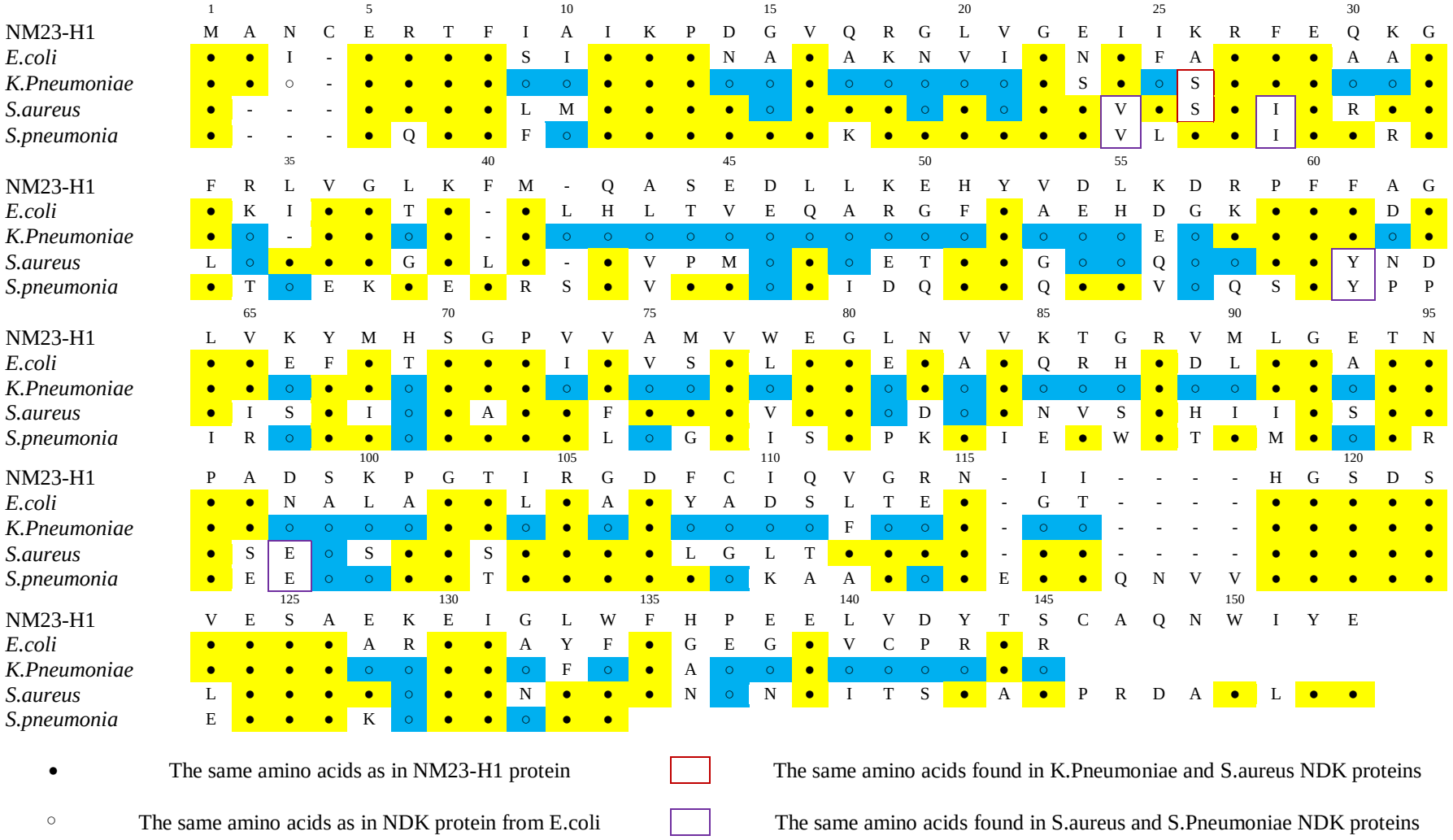
	110					115						120					125						130					135					140	
H1	M	L	G	E	T	N	P	A	D	S	-	-	-	K	P	G	T	I	R	G	D	F	C	I	Q	V	G	R	N	I	I	H	G	
H2	•	•	•	•	•	•	•	•	•	•	-	-	-	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	
H3	L	I	•	A	•	•	•	•	•	A	-	-	-	P	•	•	•	•	•	•	•	•	•	•	E	•	•	K	•	L	•	•	•	
H4	•	○	•	H	•	D	S	•	E	○	-	-	-	A	•	•	•	•	•	•	•	•	S	V	H	I	S	•	•	V	•	•	A	
H5	○	•	•	P	N	•	□	L	V	○	K	E	T	H	•	D	S	L	•	A	I	Y	G	T	D	D	L	•	•	A	L	•	•	
H6	○	M	•	◊	•	R	V	F	R	○	R	H	V	□	•	◊	◊	•	•	•	S	•	◊	L	T	◊	T	•	•	T	T	•	•	
H7	S	R	K	•	K	T	I	•	L	I	-	-	-	•	•	◊	A	•	S	K	A	G	E	•	I	E	I	I	•	K	A	G	F	
H8	E	K	I	Q	G	A	N	•	P	L	-	-	-	V	N	K	K	V	I	N	L	I	D	E	○	R	K	x	A	◊	G	E	M	
H9	T	•	A	I	I	K	•	D	A	V	-	-	-	□	H	•	#	T	D	E	◊	#	M	K	x	Q	E	A	G	F	E	I	L	
H10	x	V	F	F	R	•	C	R	•	C	-	-	-	•	C	T	L	A	C	Q	Q	•	R	□	R	◊	C	•	K	L	×	V	x	

	160					165						170					175						180					185					190	
H1	-	-	E	E	L	V	D	Y	T	S	C	A	Q	N	W	I	Y	E																
H2	-	-	•	•	•	•	•	•	K	•	•	•	H	D	•	V	•	•																
H3	-	-	D	•	•	L	C	W	E	D	S	•	G	H	•	L	•	•																
H4	-	-	S	•	•	•	S	○	A	○	G	G	•	○	S	S	I	H	P	A														
H5	V	I	V	•	P	I	P	I	G	Q	A	•	K	Δ	Y	○	N	L	H	I	M	P	T	L	L	E	G	L	T	E	L	C	K	
H6	-	F	□	•	Q	R	W	•	○	E	E	E	P	Q	L	R	C	G	◊	V	C	Y	S	P	E	G	◊	V	H	Y	V	A	G	
H7	-	⌂	H	V	D	H	Q	S	R	P	F	F	N	E	⌂	•	Q	F	I	T	T	G	P	I	I	A	M	E	I	L	R	D	D	
H8	-	-	□	•	E	S	◊	C	○	•	V	Q	E	L	◊	□	□	A	x	◊	K	◊	D	A	V	I	S	K	K	V	◊	E	◊	
H9	-	-	x	K	A	G	E	E	□	F	⌂	K	L	V	H	H	M	C	C	G	S	S	H	E	◊	#	L	T	#	T	E	G	F	
H10	-	-	G	C	F	Q	⌂	•	Y	x	⌂	L	A	F	Q	F	K	D	A	×	L	×	I	F	N	N	T	W	S	N	I	H	x	

	210					215						220					225						230					235					240	
H5	N	N	P																														
H6																																	
H7	R	I	D	A	S	A	S	I	R	A	I	F	G	T	D	G	I	R	N	A	A	H	G	P	D	S	F	A	S	A	A	R	E	
H8	V	L	T	E	E	Q	V	V	N	F	Y	S	R	I	A	D	Q	C	D	F	E	E	F	V	S	F	M	T	x	G	L	S	Y	
H9	P	◊	V	x	R	R	E	Q	-	P	E	#	L	R	#	Q	Y	G	T	E	M	P	#	N	A	V	H	G	x	R	D	x	x	
H10	A	V	×	Q	D	Y	#	P	I	×	T	T	E	E	L	K	A	V	R	V	S	T	E	A	N	R	S	I	V	P	I	#	R	

	260																																
H5																																	
H6						265						270					275						280					285					290	
H7	K	F	T	N	C	T	C	C	I	V	K	P	H	A	V	S	E	G	L	L	G	K	I	L	M	A	I	R	D	A	G	F	E	
H8	E	P	Q	T	D	x	E	P	N	E	R	S	E	D	Q	P	x	V	E	A	Q	V	T	P	G	M	M	N	K	Q	D	S	L	
H9	x	D	x	E	A	P	Q	G	E	S	S	T	Q	P	R	L	K	I	T	D	L	D												
H10	G	×	Y	#	I	A	N	A	R	K	L	I	D	E	M	V	G	K	G	F	F	L	V	Q	T	K	E	V	S	M	K	A	x	

	310					315						320					325						330					335					340	
H7	Y	E	V	Y	K	G	V	V	T	E	Y	H	D	M	V	T	E	M	Y	S	G	P	C	V	A	M	E	I	Q	Q	N	N	A	



- • The same amino acids as in NM23-H1 protein
- ○ The same amino acids as in NDK protein from E.coli
- The same amino acids found in K.Pneumoniae and S.aureus NDK proteins
- The same amino acids found in S.aureus and S.Pneumoniae NDK proteins

Figure 1.5 Alignment of *E.coli, K. pneumoniae, S.pneumonia, S.aureus* NDK protein and NM23-H1 sequences.

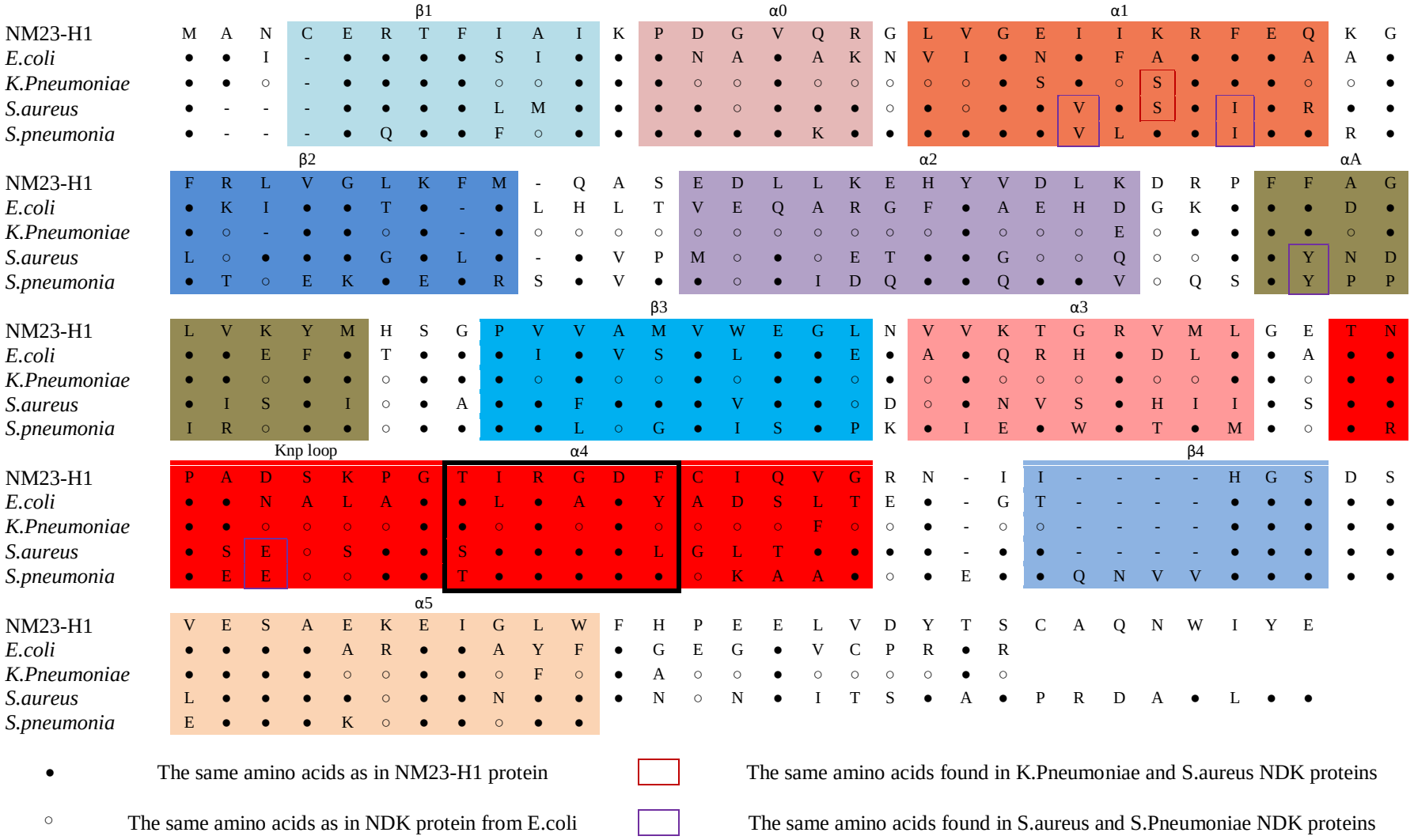


Figure 1.6 α helixes, β sheets and Knp loop matching to the *E.coli, K.pneumoniae, S.pneumonia, S.aureus* NDK proteins and NM23-H1 sequences.

1.3.1 Human NDK proteins and their function

The relationship of NM23-H1 protein expression to the suppression of metastasis firstly was reported by Steeg et al (1988). This relationship was found in such cancer types as cervical, gastric, breast, melanomas (Okabe-Kado et al., 2009) and hepatocellular carcinoma (Yamaguchi et al., 1994). Nevertheless, in some cancer types such as colon cancer (Martinez et al., 1995), neuroblastoma (Chang et al., 1994), leukaemia and other haematological malignancies (Steeg et al., 1988, Yokoyama et al., 1996, Willems et al., 1998, Niitsu et al., 2000, Okabe-Kado et al., 2002, Okabe-Kado et al., 2009) and neoplasms (Okabe-Kado et al., 2009) researchers found a relationship between an overexpression of the NM23-H1 and poor prognosis of the patients. Expression levels of NM23-H1 protein in ovarian cancer was suggested as contradictory (Okabe-Kado et al., 2009, Mandai et al., 1994).

NM23-H1 and NM23-H2 protein expressing genes are located near the breast cancer susceptibility gene 1 (BRCA1) locus at chromosome 17q21.3. BRCA is associated with development of breast and ovarian cancer (Backer et al., 1993, Saha and Robertson, 2011). Moreover, NM23-H2 was identified as a transcription factor (TF) PuF, which regulates activity of proto-oncogene c-myc (Postel et al., 1993) suggesting that changes in expression level or mutation in NM23-H2 may activate or inhibit proto-oncogene c-myc activity (Postel and Ferrone, 1994). Overexpression of c-myc plays an important role in cancer development (Miller et al., 2011), thus changes in NM23-H2 expression levels may affect cancer development. Nevertheless, Yokoyama et al (1996) observed that expression of c-myc remains unchanged in AML cells, whereas NM23-H1 and NM23-H2 are elevated.

The relationship of the other NM23 homologues with cancer development has not been reported. They play important roles in other processes and may have some effect in cell metastatic activity or its inhibition. For example, NM23-H3 is able to inhibit granulocyte

differentiation and induce their apoptosis (Mehus et al., 1999). NM23-H4 possibly is related to intramictohondrial GTP metabolism, whilst NM23-H5 may be involved in spermatogenesis (Mehus et al., 1999).

Viral pathogens such as the Epstein-Barr virus (EBV) (Murakami et al., 2009), Human Papillomavirus (HPV) (Mileo et al., 2006) and Kaposi's sarcoma associated herpesvirus (KSHV) (Qin et al., 2011) indirectly interact with NM23-H1 resulting in its suppression. EBV nuclear antigens (EBNA) 3C and 1, latency associated nuclear antigen (LANA) and E7 protein are tumour virus encoded oncoproteins for EBV, KSHV and HPV, respectively. These oncoproteins are mediators in the interaction between virus and NM23-H1. The virus triggers an expression of the oncoprotein, which binds to NM23-H1 protein, causing its translocation from cytoplasm into nucleus where it forms a stable complex with the oncoprotein. Therefore the host cell becomes transformed, resulting as a range of human cancers, for example, Burkitt's lymphoma (from EBV) (Murakami et al., 2009), cervical and ano-genital cancers (from HPV) (Mileo et al., 2006) and Kaposi's sarcoma (from KSHV) (Qin et al., 2011).

1.3.1.1 Human NM23 protein role in normal haemopoiesis and leukaemia

It is important to understand the role of human NM23 proteins in normal haemopoiesis and leukaemia. NM23-H1 and NM23-H2 proteins are expressed by normal haemopoietic and leukaemia cells. However, in leukaemia patients, the levels of both proteins are elevated. Thus, NM23 proteins may play an important role in normal haemopoiesis. Elevated NM23 in leukaemia patients may play an extra role.

Expressed NM23 proteins can be found in the cell cytoplasm, transported to the outer cell membrane where they can be secreted. The highest expression of NM23-H1 and NM23-H2 has been found in bone marrow $CD34^{+ve}$ cells. Expression levels in mature cells are lower. Moreover, levels of NM23-H1 and NM23-H2 were lineage dependent. The highest expression

levels of NM23-H1 were observed in polymorphonuclear cells, but NM23-H2 was expressed more in lymphocytes (Willems et al., 1998).

In AML patient samples elevated levels of NM23-H1 and NM23-H2 was found intracellularly (Yokoyama et al., 1996) and extracellularly (Niitsu et al., 2000). Thus, it was suggested that the elevated levels of NM23-H1 and NM23-H2 proteins are related with poor patient prognosis.

Extracellular and intracellular expression of the NM23 protein has been tested in various cell lines from different cancers. Lymphoid cell lines such as HSB2, SupT1, Jurkat (Willems et al., 1998), MOLT3, MOLT4 (Okabe-Kado et al., 2002, Willems et al., 1998), BALL-1, BALM-1, HT, DB, BALM-3, SKW-4, SU-DHL-4, U-698-M, MOLT16 and HPB-ALL (Okabe-Kado et al., 2002) did not show extracellular expression of the NM23-H1 and NM23-H2. Minor extracellular expression of NM23-H1 was observed only in Raji lymphoid cells (Okabe-Kado et al., 2002).

Extracellular expression of NM23-H1 was observed on myeloid cell lines such as KG1, U937t (Willems et al., 1998), HL-60 (Okabe-Kado et al., 2002, Willems et al., 1998), NB4, HT93, ML-1, THP-1 and U937 (Okabe-Kado et al., 2002). Expression of both of these proteins was observed as well on erythroid cell lines such as K562, KU812F, HEL and M6 (Okabe-Kado et al., 2002). Therefore, extracellular expression of NM23 is dependent on the cell type.

An important observation was the decrease of extracellular NM23 protein after K562, HEL and NB4 cell differentiation initiation using All Trans-Retinoic Acid (ATRA) (Okabe-Kado et al., 2002). Thus, extracellular NM23 levels decreases with cell maturation. Moreover, it was shown that targeting NM23 proteins with monoclonal antibody specific for NM23-H1 (H1-229, immunoglobulin class IgG_{2a}) and NM23-H2 (H2-206, immunoglobulin class IgG_{2b})

HL60 and HEL cell cytolysis can be induced. Therefore, Okabe-Kado et al (2002) suggested that NM23 proteins can be a potential target in leukaemia therapy.

1.3.1.2 Effect of additional NM23 proteins on healthy haemopoietic and leukaemia cells

NM23-M2 is able to inhibit normal mouse bone marrow growth and maturation in the presence of GM-CSF or M-CSF and to induce differentiation of myeloid leukaemia M1 cells (Okabe-Kado et al., 1988). Therefore it is also known as inhibitor factor (I-factor). It possesses 98% sequence identity with NM23-H2, 91% with NM23-M1 and 89% with NM23-H1 (Okabe-Kado et al., 1992). Thus, all NM23 family members are potentially involved in inhibition of normal haemopoietic cell growth and differentiation.

Abilities of I-factor to inhibit induced differentiation of in human cell lines such as HL-60, ML-1 and K562 and mouse cell line Friend were tested. Cell differentiation in HL-60 and ML-1 was induced using 12-O-tetradecanoylphorbol-13-acetate and retinoic acid, in K562 using hemin, 1-β-D-arabinofuranosylcytosine or sodium butyrate, but in Friend using dimethyl sulfoxide or actinomycin D. I-factor did not inhibit initiated differentiation (Okabe-Kado J. et.al., 1988). In the later study Okabe-Kado et al (1995) tested the inhibition properties of human NM23 protein family members. It was found that proteins NM23-H1, NM23-H2, NM23-M1 and NM23-M2 (I-factor) inhibit the TGF-β initiated erythroid differentiation of HEL, KU812F and K562 cell lines. Moreover, Cys^{118} mutant NM23-H2 and truncated at C terminal (has only 60 amino acids) NM23-H2 protein despite of lack of NDK activity showed similar inhibition ability as rNM23 proteins. Therefore, NM23 proteins do not require NDK activity for inhibition of cell differentiation. It was also suggested that inhibition of initiated cell differentiation is cell-line dependent (Okabe-Kado et al., 1995). Important to note that only TGF-β initiated differentiation was inhibited by I-factor. It might

be that inducing agents: hemin, 1-β-D-arabinofuranosylcytosine or sodium butyrate reduces I-factor inhibitor.

Additional NM23 protein effect was studied as well on a specific cell marker sorted cells. Normal bone marrow cells were sorted for $CD34^{+ve}/CD38^{-ve}$ and $CD34^{+ve}/CD38^{+ve}$ cells. Additional extracellular NM23-H1, NM23-H2 and NM23-H3 did not affect $CD34^{+ve}/CD38^{-ve}$ primary cell differentiation, whilst more mature $CD34^{+ve}/CD38^{+ve}$ cell differentiation was altered. In the presence of all NM23 isoforms $CD34^{+ve}/CD38^{+ve}$ cells differentiation more into BFU_E, while differentiation into CFU_M was inhibited (Willems et al., 2002). Thus, it might be that bacterial NDK proteins are also able to alter haemopoietic cell differentiation. Moreover, it might be that bacterial NDK proteins affect healthy and leukaemia cells in different ways.

Recombinant NM23-H1 and NM23-H2 proteins promote growth and survival of mononuclear AML cells. Moreover, additional rNM23-H1 causes release of GM-CSF, IL-1β, IL-6 cytokines in physiologically effective levels. Using anti-TNFα, anti-IL-1β, GM-CSF and/or anti-IL-1 antibodies alone or in varied combination, partial inhibition of AML cell survival promoted by additional rNM23-H1 could be observed. An important fact is that used antibodies inhibits p38 mitogen-activated protein kinase (MAPK) cell signaling pathway. Hence, it was proposed that through the cytokine release rNM23-H1 is able activate MAPK signaling pathway and promote AML cell survival (Okabe-Kado et al., 2009).

Lilly et al (2011) and Lilly (2012) showed that additional rNM23-H1 is able to bind to the surface of $CD34^{+ve}$ and $CD34^{lo}/CD11b^{+ve}$ cells and promote their survival. Although rNM23-H1 promoted survival of both populations, rNM23-H1 binds better to more mature cells $CD34^{lo}/CD11b^{+ve}$ then to immature $CD34^{+ve}$. Lilly et al (2011) suggested that additional rNM23-H1 indirectly promotes immature cell survival.

Lilly et al (2011) showed for the first time that rNM23-H1 indirectly promotes stem cell survival and proliferation. The authors performed an experiment where mononuclear AML cells were sorted for $CD34^{-ve}/CD117^{-ve}$ and $CD34^{+ve}/CD117^{+ve}$ cells. Then recombinant NM23-H1 was added to more mature $CD34^{-ve}/CD117^{-ve}$ cell population. $CD34^{-ve}/CD117^{-ve}$ cells incubated with additional rNM23-H1 secreted a combination of cytokines, predominantly IL1β and IL6. $CD34^{+ve}/CD117^{+ve}$ cells were resuspended in CM_{dep} collected from $CD34^{-ve}/CD117^{-ve}$ treated cell cultures. Hence, it was proposed that cytokines found in CM are promoting survival and proliferation of stem cells ($CD34^{+ve}/CD117^{+ve}$ cells).

As mentioned above, expression levels of rNM23-H1 drop with blood cell maturation. Moreover, the increased levels of rNM23-H1 were observed in AML patients. Recombinant NM23-H1 express by stem cells may bind to more mature cells and promote the release of cytokines that stimulate the proliferation of stem cells. LSCs express higher levels of rNM23-H1. The levels of cytokines released by more mature leukaemia cells may be higher than healthy cells release them. Thus, it might be that through the secretion of NM23-H1 LSCs affect more mature cells to release LSC survival promoting cytokines. We assume that rNDK proteins may act in the same way as rNM23-H1, observed by Lilly et al (2011). It might be that leukaemia patient prognosis worsens due to bacterial NDK proteins that indirectly promote leukaemia stem cell proliferation.

1.3.2 Bacterial NDK protein role in bacteria

Structural similarity between human and bacterial proteins may also mean that bacterial NDK proteins play a similar role in healthy and leukaemia patients as NM23-H1 does. Thus, in this paragraph we will discuss bacterial NDK protein role in host invasion.

1.3.2.1 Bacterial NDK secretion

Zaborina et al (1999) reported that *Pseudomonas aeruginosa* secretes enzymes involved in ATP synthesis, including NDK protein (Zaborina et al., 1999). In later studies, it was suggested that DXXX (X is predominantly hydrophobic amino acid) motif in the carboxyl-terminal region (C-terminal) of NDK protein is essential for their secretion by the type I mechanism (Kamath et al., 2000). For example, the motif in the NDK of *P.aeruginosa* it is DTEV. Secretion of NDK was observed as well in such bacteria as *Porphyromonas gingivalis* (Yilmaz et al., 2008), *Leishmania amazonensis* (Kolli et al., 2008), *M.tuberculosis* (Zhao et al., 2007), *P.aeruginosa* (Zaborina et al., 1999), *Vibrio cholera*, and *S.typhimurium* (Dar et al., 2011).

No information is available if bacteria can secrete NDK protein. There is a possibility that some bacteria do not secrete it. Moreover, the exact mechanism of NDK protein secretion in above mentioned bacteria is not determined. It remains uncertain if common bacteria such as *E.coli* (Kern et al., 2005, Gomez et al., 2003, Buncaneve et al., 2007), *Pseudomonas aeruginosa* (Kang et al., 2003), *Staphylococcus aureus* (Madani, 2000, Marty et al., 2006, Buncaneve et al., 2005) and *Streptococcus pneumonia* (Madani, 2000, Bucaneve et al., 2007) invading leukaemia patients are able to secrete NDK proteins.

1.3.2.2 Bacterial NDK effect on the host cells

Human beings are continuously exposed to harmful microbiological organisms. The largest number of bacterial species is harmless and few species are beneficial or pathogenic. Leukaemia patients lack immunity so even bacteria that are harmless to healthy humans may invade leukaemia patients (Bucaneve et al., 2007). As varied pathogenic bacteria and host interactions are extensively investigated, in this paragraph we will concentrate on pathogenic bacteria and their effects on the host cells.

The location of pathogenic bacteria in the host determines processes within bacteria and their effects on a host. Based on their location, harmful bacteria are divided into extracellular and intracellular pathogens. Pathogenic extracellular bacteria do not enter host cells; thus they usually promote host cell death. Intracellular bacteria grow and reproduce within host cells (Spooner and Yilmaz, 2012, Nazareth et al., 2007, Lau et al., 2007, Martner et al., 2008, Lowy, 2000). Such bacteria, like *Chlamydophila* spp, *Ehrlichia* spp and *Rickettsia* spp are obligate intracellular parasites (Fritsche et al., 1999), which need to prevent host cell induced apoptosis. Bacteria such as *P.gingivalis, S.typhimurium, Chlamidia trachimatis, Bacillus anthracis, Mycobacteria* spp and *L. amazonensis* are usually intracellular (Spooner and Yilmaz, 2012). On the other hand, *P.aeruginosa*, *V.cholerae* (Spooner and Yilmaz, 2012), *E.coli* (Nazareth et al., 2007), *K.pneumoniae* (Lau et al., 2007), *S.pneumonia* (Martner et al., 2008) and *S.aureus* (Lowy, 2000) are extracellular bacteria. It might be that extracellular and intracellular pathogen NDK proteins act in leukaemia patients in different ways. We will further observe how these pathogens affect non leukaemia patients, and what the role of NDK in the bacterial invasion is.

Yilmaz et al. (2008) reported that NDK of *P.gingivalis* prevents host cell apoptosis through the inhibition of $P2X_7$ receptor of gingival epithelial cells (GECs). $P2X_7$ receptor mediates immune processes such as cell motility and microbial phagocytosis (Bours et al., 2006) and the release of proinflammatory cytokines such as IL-1β and IL-18 in response to microbial invasion (Pedra et al., 2009). $P2X_7$ receptors are found in macrophages, monocytes and microglia (Lister et al., 2007). It might be that NDK of intracellular bacteria may affect leukaemia cells through the inhibition of $P2X_7$ receptors, causing them to release cytokines IL-1β and IL-18. The released cytokines observed in more mature haemapoietic cells treated with rNM23-H1 by Lilly et al (2011) may promote leukaemia stem cells to enhance survival

and proliferation of LSCs. However, *P.gingivalis* invades GECs, but not blood cells. Bacterial NDK proteins can be secreted by bacteria in this case *P.gingivalis* and transported out of host cells. Pedra et al (2009) made a conjecture that NDK protein of *P.gingivalis* may be transported from the invaded GEC to the neighboring cells. The latter suggestion, however, remains unproven.

Dar et al (2011) tested how *S.typhimurium* and *V.cholerae* bacteria secretion products containing NDK proteins, affect ATP-induced J774 mouse macrophage cell lysis. Through sets of experiments, the authors determined that secreted NDK proteins by extracellular and intracellular bacteria have similar functional properties. Secretion products of intracellular pathogen *S.typhimurium* containing either own NDK protein or NDK protein of extracellular pathogen *V.cholerae* inhibited ATP-induced J774 cell lysis. Moreover, lack of NDK in the secretion products of *S.typhimurium* did not protect J774 cells against induced cell lysis. Secretion products of the extracellular pathogen *V.cholerae* had a cytotoxic effect on ATP-non-induced and ATP-induced J774 cells and even when secretion products were collected from NDK knockdown *V.cholerae* strains.

The functional similarity of NDK proteins of extracellular and intracellular pathogens may mean that any bacterial NDK may act in a similar way in leukaemia patients. The difference between the pathogen types is that NDK proteins of intracellular pathogens require to be secreted by bacteria and transported out of the cell.

1.4 THE AIM OF THE RESEARCH

The main aim of this research is to test if *E.coli, K.pneumoniae, S.pneumoniae* and *S.aureus* rNDK proteins promote the survival of leukaemia stem cells in the same way as it was shown by Lilly et al (2011) using rNM23-H1. We hypothesized that when bacteria invade leukaemia patients secreted bacterial NDK proteins indirectly promote survival and proliferation of leukaemia stem cell through binding to more mature cells.

2 METHODS

2.1 PRODUCTION OF HIS-TAGGED RECOMBINANT NDK (rNDK) PROTEINS

2.1.1 rNDK protein production system

Glycerol stocks (available in our laboratory) of *E.coli* BL21 (DE3) (Bioline, London, UK) cultures containing the *E.coli*, *S.aureus*, *S.pneumoniae* or *K.pneumoniae* rNDK gene coded in the pET-15b expression vector (Novagen, Merk Chemicals Ltd, Nottingham, UK) were used. The pET-15b expression vector contains an ampicillin resistance marker, T7 promoter, Lac operator and 6×His-tag sequence coding gene (Figure 2.1). NDK protein coding gene is ligated in-between Nde1 and BamH1 restriction sites of pET-15b expression vector (Figure 2.1)

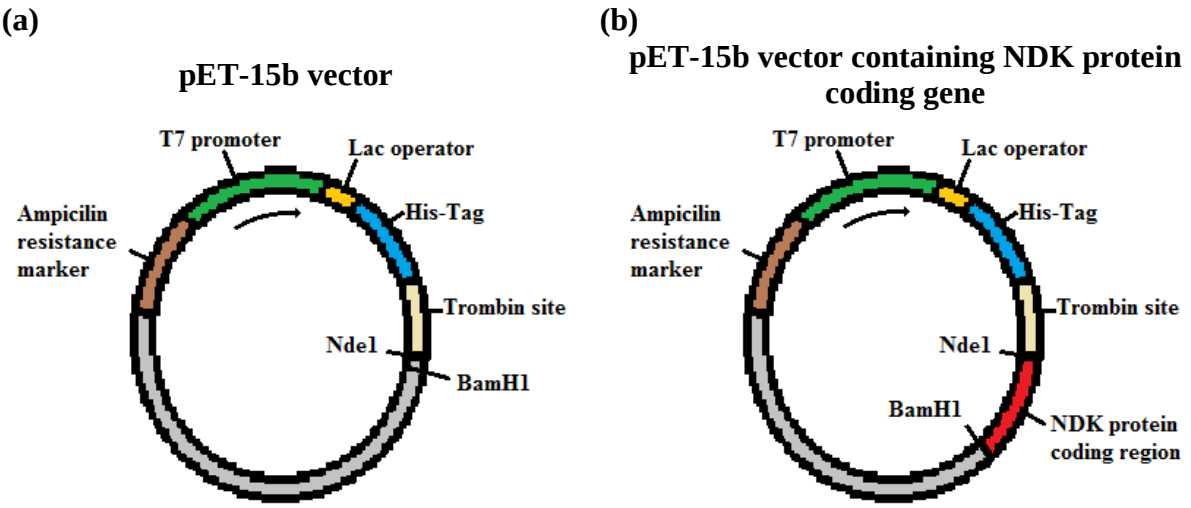


Figure 2.1 Commercial pET-15b vector and pET-15b vector containing NDK protein coding gene. (a) pET-15b vector purchased from Novagen. (b) pET-15b vector containing NDK protein coding gene.

2.1.2 rNDK protein expression

BL21 (DE3) bacteria glycerol stocks were streaked out on LB agar plates supplemented with 100 µg/ml ampicillin (see Appendix) and grown overnight at 37°C. Three different cultures of BL21 (DE3) for each rNDK were picked and inoculated in 5 ml LB media supplemented with 100 µg/ml ampicillin (see Appendix) and incubated overnight at 37°C with shaking at 220 rpm. The integrity of overnight cultures was checked by colony PCR. Overnight cultures were diluted 1:50 in 150 ml LB media supplemented with 100 µg/ml ampicillin and incubated for 4 hours at 37°C with shaking at 220 rpm. After 4 hours 1mM an isopropyl β-D-1-thiogalactopyranoside (IPTG) (Sigma-Aldrich Ltd., Gillingham, UK) was added to induce rNDK protein expression. IPTG induced cultures were left incubating for 24 hours at 37°C with shaking at 220 rpm.

2.1.3 rNDK protein purification

After 24 hour incubation IPTG induced cultures were centrifuged at 6000×g at 4°C for 10 minutes. Supernatant was removed, and wet mass weighed. To lyse bacteria pellets were resuspended in 5 ml of BugBuster™ (Novagen) per 1 g of wet weight and 1 µl/ml benzonase nuclease (Novagen). Cell lysis was performed at room temperature for 40 minutes, following incubation at 4°C for 20 minutes. Lysate was centrifuged at 14000 rpm for 20 min. Supernatant was collected, whilst pellets were discarded. rNDK proteins were purified from supernatant using the His-Bind Purification kit (Novagen) following manufacturer's instructions. Each column was filled with 2 ml Ni-NTA His-Bind resin, where resin settled bed volume was 1 ml. When liquid stopped dripping, equilibration and charge of the resin was performed following 3 steps:

1) 3 ml sterile dH_2O
2) 5 ml 1×Charge Buffer

3) 3 ml 1×Binding Buffer

Supernatants containing rNDK proteins were loaded onto the columns and allowed to drip through. Columns were washed with 10 ml 1×Binding Buffer, then with 6 ml 1×Wash buffer. Proteins were eluted with 6 ml 1×Elute Buffer and collected in 6 separate 1 ml fractions.

Protein elution was verified by sodium dodecyl sulphate (SDS) polyacrylamide gel staining with Coomassie blue and Western blot analysis. rNDK protein concentration was determined for the first three fractions. Only pure fractions were used for further experiments.

2.2 MEASURING SECRETION OF rNDK PROTEINS FROM BL21 (DE3) BACTERIA

Overnight cultures (described above) were used for the time course experiment. Overnight cultures were diluted 1:50 in 10 ml LB media supplemented with 100 µg/ml ampicillin and incubated for 4 hours at 37°C with shaking at 220 rpm. After 4 hours, 1mM IPTG was added to induce expression of rNDK proteins. IPTG non-induced cultures were used as controls. IPTG induced and non-induced cultures were left incubating at 37°C with shaking at 220 rpm for the next 6 days. Zero time point samples (1 ml) were collected shortly before IPTG induction. The next samples were collected as 1 ml every 24 hours.

Samples were centrifuged at 14,000 rpm for 5 minutes and pellets were discarded. Supernatants were separated by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) on 12.5% gels and stained with Coomassie blue. Secreted rNDK proteins were verified by calculation of their masses and comparison to the expected masses.

2.3 CELL PURIFICATION AND CELL CULTURE

Umbilical cord blood (CB) samples were provided by tissue bank under ethical approval following informed consent. CB samples were received in plastic bags supplemented with anti-coagulants. Cell purification was performed as described in Figure 2.2.

2.3.1 The purification of mononuclear cells from umbilical cord blood (CB)

Mononuclear cell separation was performed using 50 ml tubes and spin filter tubes (PPA, London, UK). 15 ml Ficoll-Paque^{TM} PLUS (G.E.Healthcare, Amersham, UK) was pipette on the top of filter in spin filter tubes. To move the Ficoll-Paque^{TM} PLUS through the filter, tubes were centrifuged at 1000×g for 1 minute. Blood samples were diluted 1:3 in RPMI 1640 (Gibco, Life Technology, Paisley, UK) supplemented with 1% v/v penicillin and streptomycin, and 1% v/v ITS (RPMI 1640 p/s ITS+) and layered on the top of filter. The samples were centrifuged at 1000×g for 10 minutes. The mononuclear cell layer is found just above the filter in spin filter tubes. Plasma was removed and mononuclear cells were collected into a clean 50 ml tube. To wash mononuclear cells, they were resuspended 1:5 in RPMI 1640 p/s ITS+ and centrifuged at 374×g for 10 minutes. Supernatant was removed. Mononuclear cells were resuspended in 5 ml RPMI 1640 p/s ITS+ and centrifuged at 249×g for 10 minutes. The last wash step was repeated twice. After the last wash, cells were resuspended in 20 ml RPMI 1640 p/s ITS+ and the number of mononuclear cells was counted using a haemocytometer. Samples were centrifuged at 1500 rpm for 5 minutes and supernatant was removed.

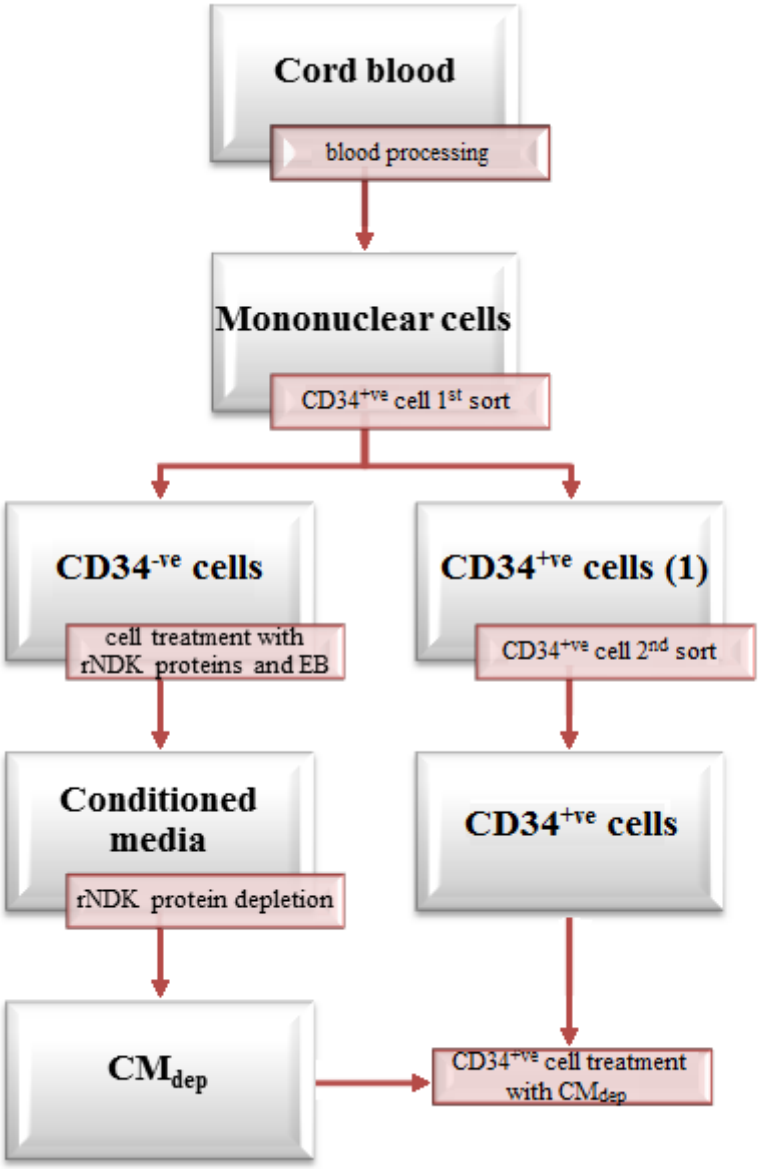


Figure 2.2 Cord blood sample processing and CD34^{+ve} sorting, where CM_{dep} is rNDK depleted conditioned media.

2.3.2 Sorting of CD34^{+ve} cells from primary mononuclear cells

Mononuclear cells were washed in 5 ml cold (6°C - 12°C) magnetic activated cell sorting (MACS) buffer (see Appendix) and centrifuged at 374×g for 5 minutes. CD34^{+ve} cells were sorted from mononuclear cells using indirect CD34 Microbead kit (Miltenyi Biotec, Surrey, UK) following manufacturer's instructions. Briefly, every 10^8 mononuclear cells were resuspended in 300 µl MACS buffer, 100 µl FcR Blocking Reagent and 100 µl CD34 Hapten-Antibody were added, then mixed and incubated for 15 minutes at 6°C - 12°C. Mononuclear cells incubated with CD34 Hapten-Antibody were washed with 10 ml MACS buffer and centrifuged at 1500 rpm for 5 minutes. Supernatant was removed and cells were resuspended in 400 µl MACS buffer. Anti-Hapten MicroBeads (100 µl) were added, mixed and incubated at 6°C - 12°C for 15 minutes. After incubation with Anti-Hapten MicroBeads mononuclear cells were washed with 10 ml MACS buffer and centrifuged at 1500 rpm for 5 minutes. Supernatant was removed and cells were resuspended in 500 µl MACS buffer.

CD34^{+ve} cells were sorted using MS column for < 2×10^8 mononuclear cells or LS Column for < 2×10^9 mononuclear cells. Briefly, MS column was placed into a magnet and washed with 500 µl MACs buffer. Mononuclear cells stained with CD34 Hapten-Antibody and Anti-Hapten MicroBeads were applied to the MS column and washed with 500 µl MACS buffer three times. CD34^{-ve} cells were collected in the unbound fraction. The column was removed from the magnet and placed on the top of a clean tube. CD34^{+ve} cells were eluted with 1 ml MACS buffer by using the plunger supplied with the column. To increase purity of eluted CD34^{+ve} cells collected CD34^{+ve} cell fraction was applied to a new MS column. MS column was prepared as described above. CD34^{+ve} cell were eluted with 500 µl MACS buffer.

CD34^{+ve} and CD34^{-ve} eluted fractions were centrifuged at 1500 rpm for 5 minutes and supernatant was removed. CD34^{-ve} cells were resuspended in RPMI 1640 p/s ITS+ as 10^6

cells per 1 ml. $CD34^{+ve}$ cells were resuspended in 2 ml RPMI 1640 p/s ITS+ and kept at 4°C overnight.

The percentage of $CD34^{+ve}$ cells in mononuclear cell fraction (pre $CD34^{+ve}$ sort sample) and the purity of eluted $CD34^{+ve}$ cells were analyzed by flow cytometer.

2.3.3 $CD34^{-ve}$ cell treatments

Resuspended $CD34^{-ve}$ cells in RPMI 1640 p/s ITS+ (500 µl) were seeded as 0.5×10^6 cells per well into a 48 well plate. For each treatment with each rNDK protein or elution buffer (EB) (Novagen) $CD34^{-ve}$ cells were seeded into two wells. Seeded $CD34^{-ve}$ cultures were treated with 1 µg/ml or 2 µg/ml rNDK proteins or EB as a negative control. $CD34^{-ve}$ cell cultures of CB12 sample were supplemented with 1.25 µg/ml polymyxin B sulfate salt (PMB) (Sigma). One of the $CD34^{-ve}$ cultures for the same treatment was incubated for 20 hours, the other for seven days. From samples incubated for 20 hours conditioned media (CM) were removed and from collected CM rNDK were depleted. Samples of $CD34^{-ve}$ incubated for seven days were collected for flow cytometry analysis and Jenner Giemsa staining.

2.3.4 $CD34^{+ve}$ cell treatments

Only twice sorted $CD34^{+ve}$ cells were used for treatments. Overnight stored $CD34^{+ve}$ cells were count using a haemocytometer on the second day after the CD34 sorting. $CD34^{+ve}$ cells were spin at 2000 rpm for 5 minutes and supernatant was removed. $CD34^{+ve}$ cells were resuspended as 1×10^6 cells per 1 ml in rNDK protein depleted CM (CM_{dep}), seeded as 0.15×10^6 or 0.1×10^6 cells per well into a 96 well plate and incubated for 7 days. On the day 7 samples of treated $CD34^{+ve}$ cells were collected for flow cytometry analysis and Jenner Giemsa staining.

2.4 COLONY PCR

The integrity of bacterial colonies was verified for the overnight cultures by colony PCR amplification of the rNDK gene cloned in the pET-15b expression vector using bacterial NDK specific primers. PCR using NM23-H1 specific primer was run on all colonies.

PCR mix (see Appendix) was prepared. Overnight culture (5 µl) was diluted in 5 µl dH_2O and by heating the mix at 90°C for 10 minutes cDNA was released. 1×PCR mix, 2 µl 10 µM forward primer (see Appendix) and 2 µl 10 µM reverse primer (see Appendix) was added to prepared cDNA. cDNAs were amplified in the PCR cycle:

94°C for 10 minutes

94°C for 15 seconds
56°C for 15 seconds } 40 cycles
72°C for 40 seconds

72°C for 10 minutes

Each PCR product (10 µl) was mixed with 10 µl 5×DNA gel loading buffer (Bioline, London, UK). Samples were load on 1% agarose gel (see Appendix) containing 0.3 µg/ml ethidium bromide (Sigma). Agarose gel electrophoreses was performed in 1×TBE buffer (see Appendix) at 125 V for 1 hour.

2.5 THE DETERMINATION OF rNDK PROTEIN CONCENTRATION

The determination of rNDK protein concentration was performed using DC Protein Assay kit (BioRad, Hemel Hempstead, UK) following manufacturer's instructions. Briefly, reagent A' was prepared by mixing 20 µl reagent S with 1 ml reagent A. Each purified rNDK protein fraction (2 µl) and 2 µl standard at varied concentrations (bovine serum albumin (BSA) diluted in RIPA buffer (see Appendix) as 0, 0.6, 1.25, 2.5, 5.0 and 10.0 mg/ml) were pipetted into a 98 well plate. Reagent A' (25 µl) and 200 µl reagent B was added to each well

and left incubating for 15 minutes. The intensity of each well was read using plate reader. Measured standard values were used to draw a standard curve. Concentrations of rNDK proteins were calculated using standard curve.

2.6 PROTEIN DEPLETION FROM CM

Treated $CD34^{-ve}$ cultures were collected in ependorfs as 500 µl after 20 hours incubation at 37ºC. Collected cultures were spin at 15,000 rpm for 15 seconds and pellet free CMs were collected in new ependorfs. HIS-Select™ Nickel affinity gel (Sigma) is used as 100 µl per 1 ml of CM to deplete rNDK protein. For example, for rNDK protein depletion from 3 ml of CM (6 cultures each 500 µl) 300 µl HIS-Select™ Nickel affinity gel is needed. HIS-Select™ Nickel affinity gel (300 µl) was placed in ependorf, spin at 15,000 rpm and supernatant was removed. HIS-Select™ Nickel affinity gel was washed with 1 ml dH_2O and was spin at 15,000 rpm. Supernatant was removed. Then HIS-Select™ Nickel affinity gel was resuspend in 1 ml equilibration buffer (see Apendix) and split equally into 6 ependorfs. To each ependorf with HIS-Select™ Nickel affinity gel 500 µl CM was added and incubated at 4°C with shaking for 20 minutes. After incubation CM and gel were spin at 15,000 rpm for 2.5 minutes. CM_{dep} was removed and filter sterilized.

2.7 SDS–PAGE

2.7.1 Sample preparation

Eluted rNDK protein fractions (Paragraph 2.1.3), supernatants collected from bacteria cultures (Paragraph 2.2) CM and CM_{dep} (Paragraph 2.3.3) were diluted 3:4 in 4×gel loading buffer (GLB) (see Appendix). For Western blot analysis the second eluted fractions of rNDK (Paragraph 2.1.3) were diluted 1:10 in 1×GLB and as 1mg/ml.

Prepared samples were boiled for 10 minutes. Each sample (10 μl) was load on to SDS-polyacrylamide gel. Coomassie blue staining was performed to check purity of eluted rNDK fractions. Western blot was performed to check rNDK depletion in CM and to check the second eluted fraction for rNDK proteins.

2.7.2 SDS – polyacrylamide gel preparation

Recombinant NDK proteins were separated by SDS–PAGE on 1.5 mm 12.5% SDS–polyacrylamide gels. SDS–polyacrylamide gel consist of two layers: resolving gel (see Appendix) on the bottom and stacking gel (see Appendix) on the top. Resolving gel (8 ml) was pipetted in-between glass plates. Water (1 ml) was pipetted on the top of the gel and resolving gel was left to polymerize for 30 minutes. Water layer was removed and prepared stacking gel was pipetted on the top of resolving gel layer. A 10 well comb was inserted and stacking gel was left to polymerize for next 30 minutes.

Two prepared SDS–polyacrylamide gels were placed into a box. Running buffer (see appendix) was added in between two gels and into the box to cover the bottom of both gels. Prepared samples and protein ladder were loaded on SDS–polyacrylamide gel. Gels were electrophoresed at 120 V for approximately 1 hour.

2.8 COOMASSIE BLUE STAINING

After SDS–PAGE gels were stained with Coomassie stain (see Appendix) with shaking at 20 rpm for 2 hours. Coomassie stain was removed; gels were placed in Destain (see Appendix) and shaken at 20 rpm until gels are transparent. Stained proteins on SDS–polyacrylamide gels were analyzed using a GeneFlash transluminator (Geneflow, Staffordshire, UK).

2.9 WESTERN BLOT ANALYSIS

Polyvinilidene difluoride (PVDF) membrane (Millipore Corp, Bedford, MA, USA) was soaked in 100% methanol for 10 seconds, and then washed in transfer buffer (see Appendix). Filter papers (12 pieces per gel) were soaked in transfer buffer. Protein transfer from SDS-polyacrylamide gel to PVDF membrane was set up as follows: 6 pieces of filter paper on the bottom, PVDF membrane, gel and 6 pieces of filter paper on the top. Protein transfer was performed at 25 V for 1 hour.

After transfer PVDF membrane was washed in TBS-T (see Appendix) with shaking at 60 rpm for 5 minutes, three times. NM23-H1 (C-20) rabbit polyclonal primary antibodies (Santa Cruzm, Wembley, UK) were diluted 1:5000 in 10 ml 5% Blotto, but His-probe (H-15) rabbit polyclonal primary antibodies (Santa Cruz) were diluted 1:2500 in 10 ml 5% Blotto. Anti-rabbit secondary antibodies (Sigma) were diluted 1:1000 in 10 ml 5 % Blotto. To prevent non-specific antibody binding PVDF membrane was soaked in 30 ml 5% Blotto (see Appendix) with shaking at 20 rpm for 30 minutes. Blotto was poured, off and membrane was incubated with diluted primary antibodies at 4°C with shaking at 20 rpm, overnight. The excess of primary antibodies was removed by washing PVDF membrane in TBS-T with shaking at 60 rpm for 10 minutes, three times. Then PVDF membrane was incubated with diluted secondary antibodies with shaking at 20 rpm, for 1 hour. The excess of secondary antibodies was removed by washing PVDF membrane in TBS-T with shaking at 60 rpm for 10 minutes, three times.

SuperSignal West Pico Chemiluminiscent Substrate or SuperSignal West Dura Chemiluminiscent Substrate (Thermo Scientific, Waltham, UK) was used. Stable peroxide solution and Luminal/Enhancer solution were mixed as 1:1 in total volume 2 ml for 1 membrane. PVDF membrane was incubated in 2 ml of the prepared solution for 5 minutes.

The excess of solution was drained off. PVDF membrane was wrapped in cling film and taped down into a film cassette. Prepared membrane was exposed to a photographic film (Kodak, Rochester, NY, USA) for a varied time and developed using AGFA Curix 60 developer (AGFA, Mortsel, Belgium).

2.10 rNDK PROTEIN MASS CALCULATION

Expected protein masses were calculated from rNDK protein sequences. Secreted rNDK protein masses were calculated from photographs of SDS-polyacrylamide gels stained with Coomassie blue. Firstly, a curve for protein ladder with respect to protein run distance was drawn. Mass of each secreted rNDK protein was determined from the equation of the trend line for the drawn curve.

2.11 FLOW CYTOMETRY

2.11.1 $CD34^{+ve}$ cell staining

$CD34^{+ve}$ pre-sorted cells, $CD34^{+ve}$ cells after first and second sort and $CD34^{+ve}$ cells incubated in CM were collected as 50 - 100 µl. To the collected samples 5 µl FcR blocking reagent (Miltenyi Biotech) and 2 µl CD34 FITC (BD Bioscience, Oxford, UK) or 1 µl CD34 APC antibody (BD Bioscience) were added. Control samples were stained with the same amount of FITC (BD Bioscience) or APC isotype control antibody (BD Bioscience). All samples were incubated for 10 minutes, washed with 2 ml PBS (PAA Laboratories, Gmbh, Yevil Sommerst, UK) then spin at 1500 rpm for 5 min. Supernatant was removed, and cells were resuspended in 200 µl FACS fix (see Appendix). Microbeads (CytoCount™, Alere Ltd., Stockport, UK) (10 µl) were added to the samples. Prepared samples were analyzed using BD LSRII machine with FACS DIVA software (BD Bioscience), collecting 100,000 viable cells for $CD34^{+ve}$ pre-sorted samples and 10,000 viable cells for $CD34^{+ve}$ sorted samples.

2.11.2 CD34^{-ve} cell staining

Treated CD34^{-ve} cells were collected as 100 - 200 µl. To the collected samples 5 µl FcR blocking reagent and 1 µl CD34 APC and 2 µl CD11b PE* antibodies (BD Bioscience) were added. Control samples were stained with the same amount of APC and PE* isotype control antibodies (BD Bioscience). Samples were incubated for 10 minutes, washed in 2 ml PBS, then spin at 1500 rpm for 5 min. Supernatant was removed, and cells were resuspended in 200 µl FACS fix. Prepared samples were analyzed using BD LSRII machine with FACS DIVA software, collecting 20,000 viable cells.

2.12 JENNER GIEMSA STAINING

Giemsa Buffer (see Appendix) was diluted 1:25 in dH_2O. Jenner stain (RAYMOND A LAMB Ltd., Eastbourne, UK) was diluted 1:3 in diluted Giemsa buffer. Giemsa's staining solution (BDH, VWR International Ltd., Poole, UK) was diluted 1:20 in diluted Giemsa buffer. To produce slides 75µl of cell culture were cytospined at 500 rpm for 3 min using Cytospin 3 (Shannon, UK). Prepared slides were air dried. Cells on the slides were fixed in methanol for 5 minutes and air dried. Slides were stained in diluted Jenner stain for 5 minutes then washed with water. Then slides were stained in dilute Giemsa stain for 10 minutes. Slides were washed, and air dried. Stained slides were mounted with DePex mounting medium (BDH, VWR International Ltd., Poole, UK). Cells were visualized using microscope. Photographs were taken at ×400 magnification.

3 RESULTS

3.1 PRODUCTION OF HIS-TAGGED rNDK PROTEINS

Glycerol stocks of BL21 (DE3) with the pET-15b expression vector containing NDK coding gene were streaked out and regrown on LB agar plates. Single colonies were picked and cultured overnight. Overnight cultures were diluted, and rNDK protein expression was induced with IPTG as described in materials and methods.

3.1.1 Integrity of the rNDK protein producing system

The integrity of bacterial colonies was verified for the overnight cultures by colony PCR amplification of the NDK gene cloned in the pET-15b expression vector using bacterial NDK specific primers. PCR using NM23-H1 specific primers was run on all colonies. PCR products were separated by agarose gels electrophoresis. The correct sized PCR products for each bacterial NDK strain and no NM23-H1 contamination were observed for all BL21 (DE3) colonies (Data not shown due to the suitable image of the gel was not captured) indicating that the rNDK colonies contained the correct plasmid sequence.

3.1.2 Purification of rNDK proteins

Recombinant NDK proteins were purified using Ni-NTA His-Bind resin from the IPTG induced cultures as described in materials and methods. Briefly, BL21 (DE3) rNDK protein producing cells were lysed, and pellets were discarded. Total lysate was added to the Ni-NTA His-Bind resin column, and unbound fractions were collected. rNDK proteins were eluted in 6 fractions as 1 ml. Purified proteins were separated by SDS-PAGE on 12.5% gels and stained with Coomassie blue (Figure 3.1). *E.coli* and *S.pneumoniae* rNDK proteins were eluted in the second and third fractions. *S.aureus* rNDK and rNM23-H1 proteins were eluted in the first three fractions, whilst *K.pneumoniae* rNDK was eluted in the first four fractions. For the first three fractions protein concentration was determined (Figure 3.2). The highest concentration of rNDK proteins is in the second eluted fraction. Therefore, the second fractions were diluted in RPMI 1640 p/s ITS+ to a final concentration of 100 µg/ml.

Verification of rNDK proteins in the second eluted fractions was performed using Western blot. There are no commercially available antibodies for bacterial NDK proteins so we tested anti-NM23 antibodies for their cross-reactivity against bacterial NDK proteins. In the first Western blot analysis NM23-H1 antibody was used (Figure 3.3). Bands were observed for rNM23-H1, *S.aureus* rNDK and their dimers. NM23-H1 antibody is rNM23-H1 specific, thus when protein fractions are diluted as 1:10 in 1×GLB and loaded on a SDS-polyacrylamide gel, photo bleaching of rNM23-H1 was observed. Bands of *K.pneumoniae* and *E.coli* rNDK usually were observed, when membrane was incubated in strong ECL and exposed to a film for a longer time. No band for *S.pneumoniae* rNDK was observed. Therefore, alignment of bacterial NDK protein sequences with NM23-H1 was produced (Figure 3.4).

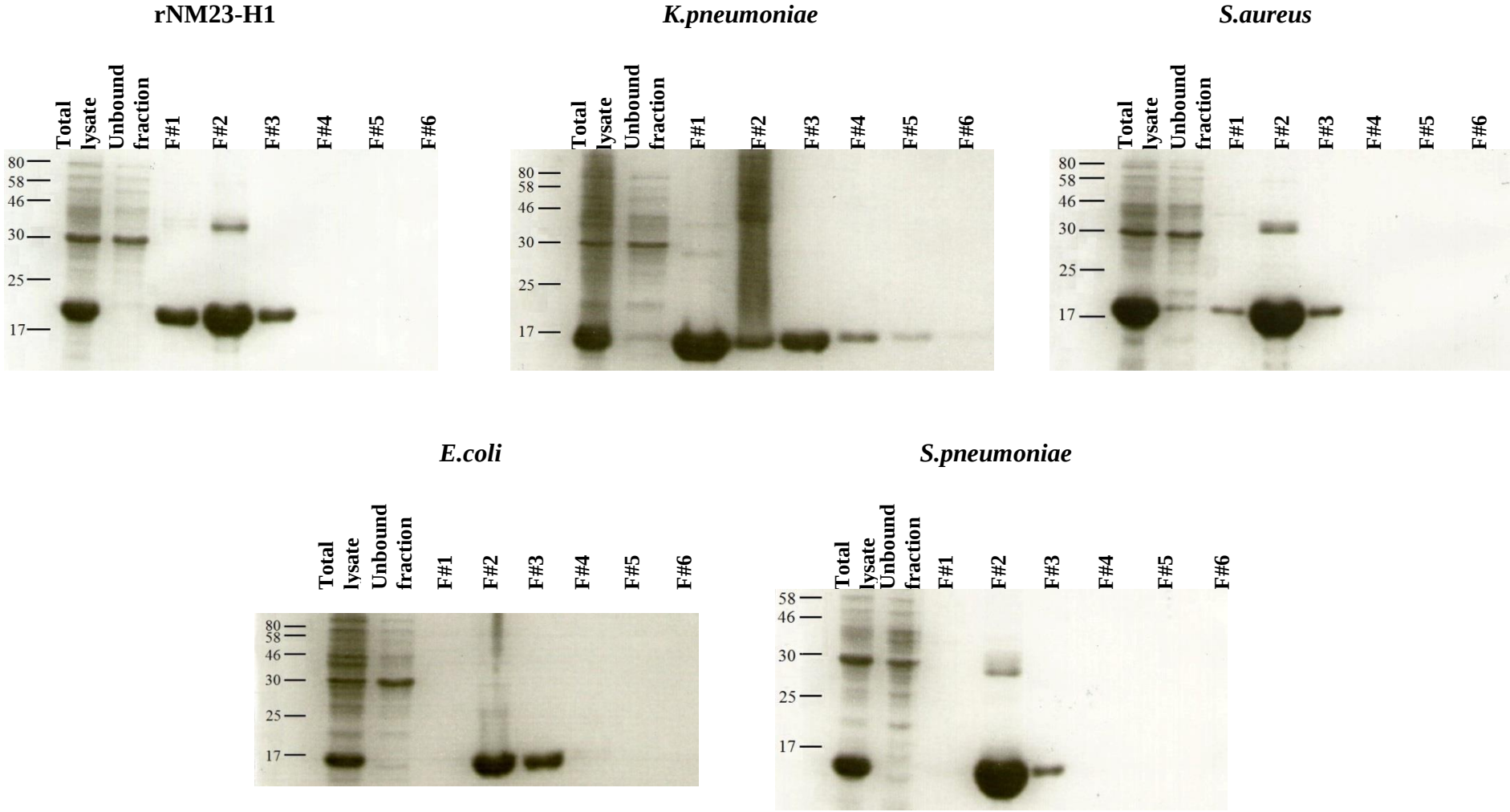


Figure 3.1 Purification of rNDK proteins. Coomassie blue stained SDS-polyacrylamide gels. For each rNDK protein samples of total cell lysate, unbound fraction and six eluted rNDK protein fractions (F#1 –F#6) were loaded on a gel.

NM23-H1 antibody binds to the 20 amino acid long C-terminal NM23-H1 sequence. Sequence similarities between NM23-H1 and *E.coli, K.Pneumoniae, S.pneumoniae, S.aureus* NDK proteins are 15%, 15%, 5% and 8 amino acids 40% (Figure 3.3). This explains observed band intensity and absence of *S.pneumoniae* rNDK band. To overcome the problem, in further Western blot analysis His-probe antibody was used.

The second eluted fractions of rNDK proteins were diluted as 1 mg/ml and loaded as 10 µl on a SDS-polyacrylamide gel. His-probe antibody binding to all rNDK proteins was observed (Figure 3.3). Weak binding of His-probe antibody to *S.pneumonia* rNDK was observed. His-probe antibody binds stronger to *E.coli, K.Pneumoniae, S.aureus* rNDK and rNM23-H1 proteins. For rNM23-H1 and *E.coli, S.aureus* rNDK were observed two bands – mono and dimer form of rNDK.

Protein	**Concentration of F#1, mg/ml**	**Concentration of F#2, mg/ml**	**Concentration of F#3, mg/ml**
rNM23-H1	0.71	2.57	0.69
E.coli rNDK	0.00	2.30	0.92
K.pneumonia rNDK	2.35	4.15	2.47
S.pneumoniae rNDK	0.49	4.30	2.00
S.aureus rNDK	0.27	3.46	1.13

Figure 3.2 Concentration of rNDK proteins in the first three fractions, where F#1, F#2 and F#3 are fractions 1-3.

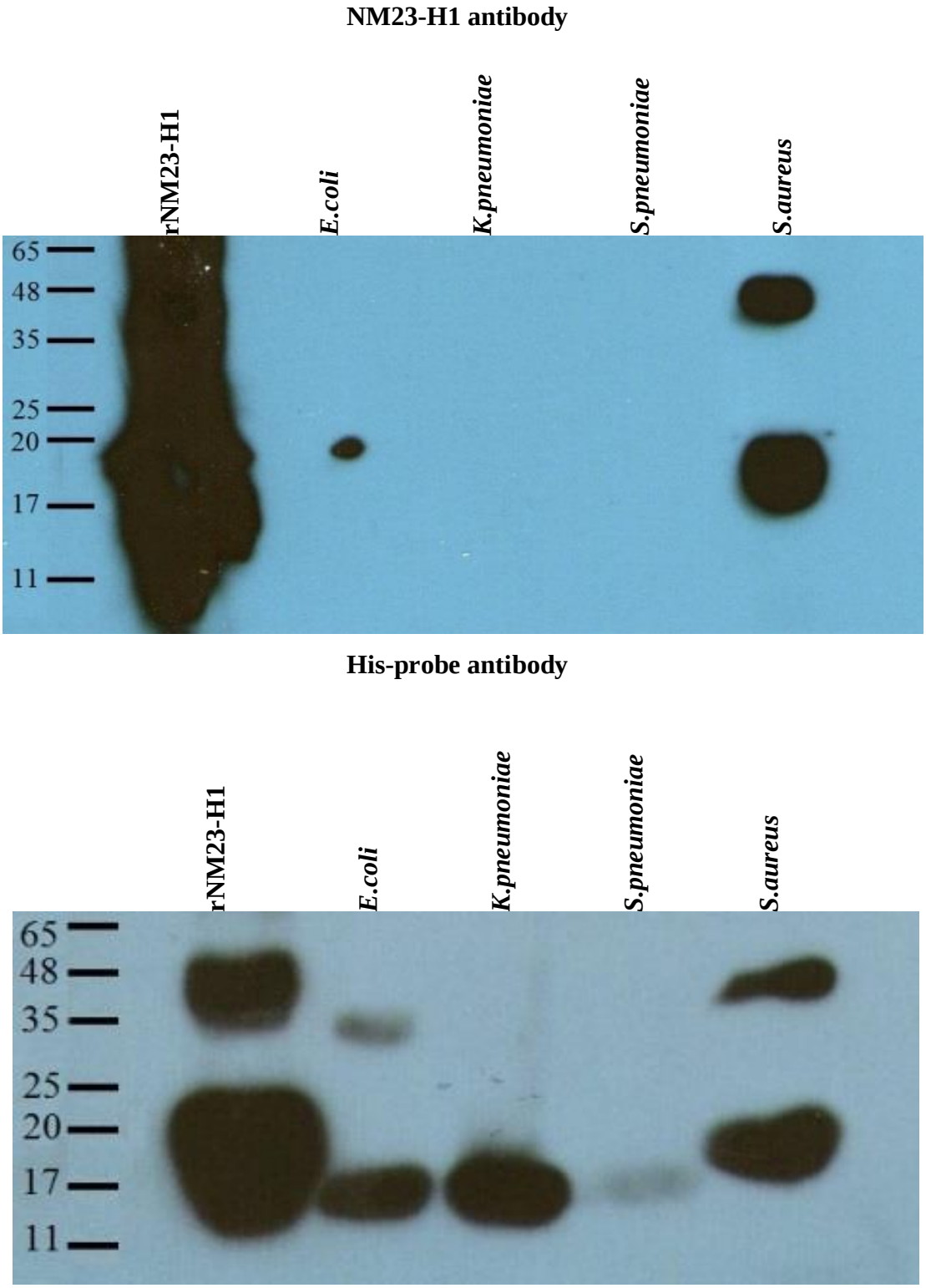


Figure 3.3 Purification of rNDK proteins. Verification of rNDK proteins using Western blot. The second eluted fraction of each rNDK protein was diluted 1:10 in 1×GLB (for PVDF membrane incubation in NM23-H1 antibody) and 1mg/ml in GLB (for PVDF membrane incubation in His-probe antibody) and loaded on SDS-polyacrylamide gel. Proteins separated by SDS-PAGE were transferred from SDS-polyacrylamide gel onto PVDF membrane and membrane was incubated in **N**M23-H1 or His-probe antibody.

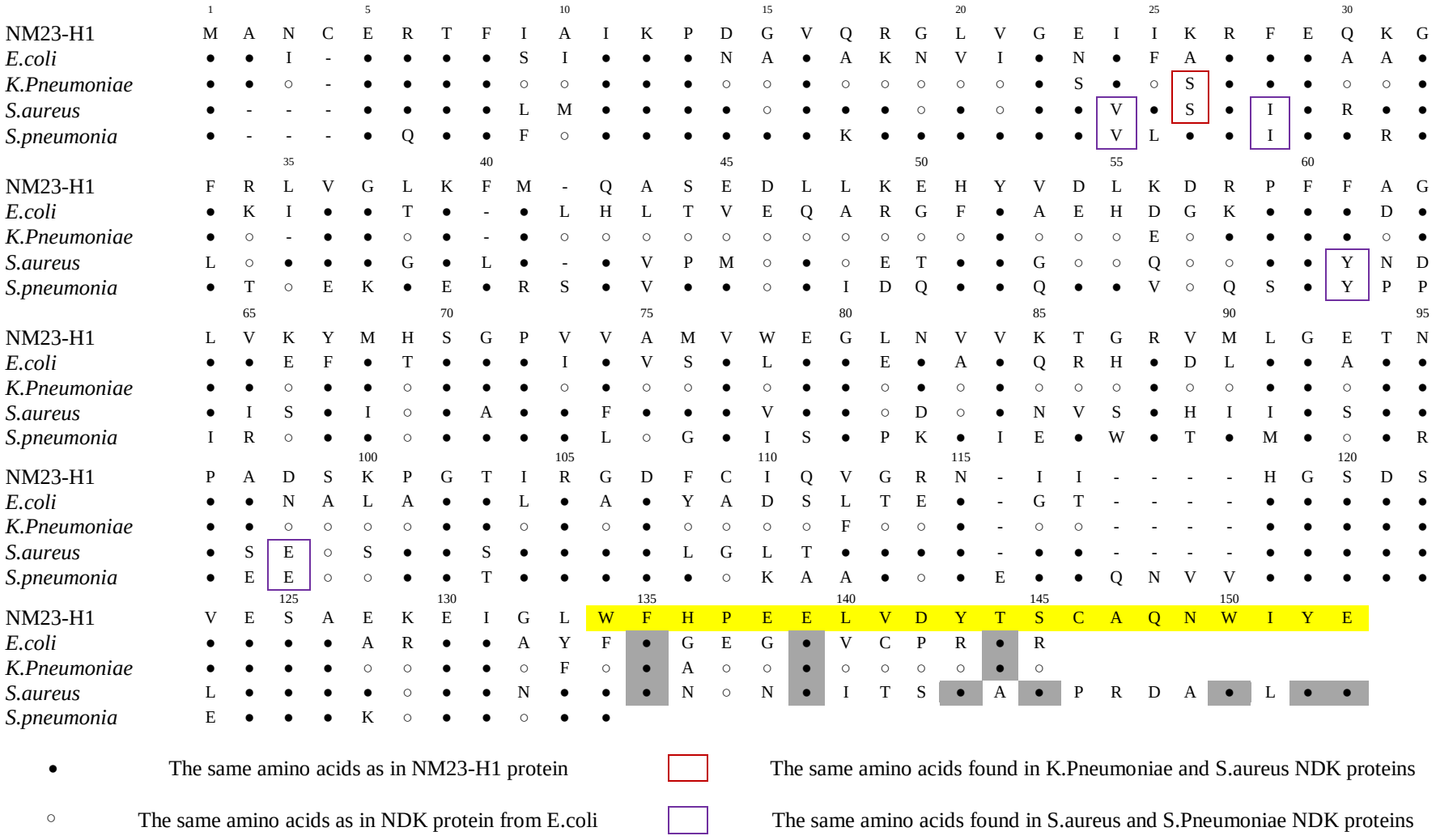


Figure 3.4 Alignment of NM23-H1 and *E.coli, K.Pneumoniae, S.pneumoniae, S.aureus* NDK protein sequences. NM23-H1 protein C-20 amino acid sequence with which NM23-H1 protein binds to NM23-H1 antibody is shown in yellow. Bacterial NDK protein similarity with NM23-H1 protein C-20 sequence is shown in grey.

3.2 BL21 (DE3) BACTERIA SECRETE rNDK PROTEINS

To check if BL21 (DE3) bacteria secrete rNDK proteins overnight cultures were diluted and rNDK protein expression was induced by IPTG. Samples were collected shortly before IPTG induction (zero time point) and every 24 hours for six days. IPTG non-induced culture was monitored as a control. Secreted proteins were separated by SDS-PAGE on 12.5% gels and stained with Coomassie blue (Figure 3.5).

For each rNDK protein curves of protein ladder mass with respect to their run distance were drawn. Secreted protein masses were determined from the trend line equation of the curve (Figure 3.6). Verification of rNDK proteins was performed by calculation of masses for observed proteins on Coomassie blue stained SDS-polyacrylamide gels and comparison to the predicted masses. Calculated masses were found as similar to predicted ones.

No rNDK secretion in IPTG non-induced cultures was observed (Data not shown). Secretion of rNDK proteins in IPTG induced cultures was observed from the first day (Figure 3.5). The intensity of observed rNDK protein bands allow to suggest that secretion of *K.pneumoniae, S.aureus* and *S.pneumoniae* rNDK stays unchanged for 6 days, while secretion of rNM23-H1 and *E.coli* rNDK decreases. Moreover, *E.coli* rNDK secretion is weak and is observed only for the first three days.

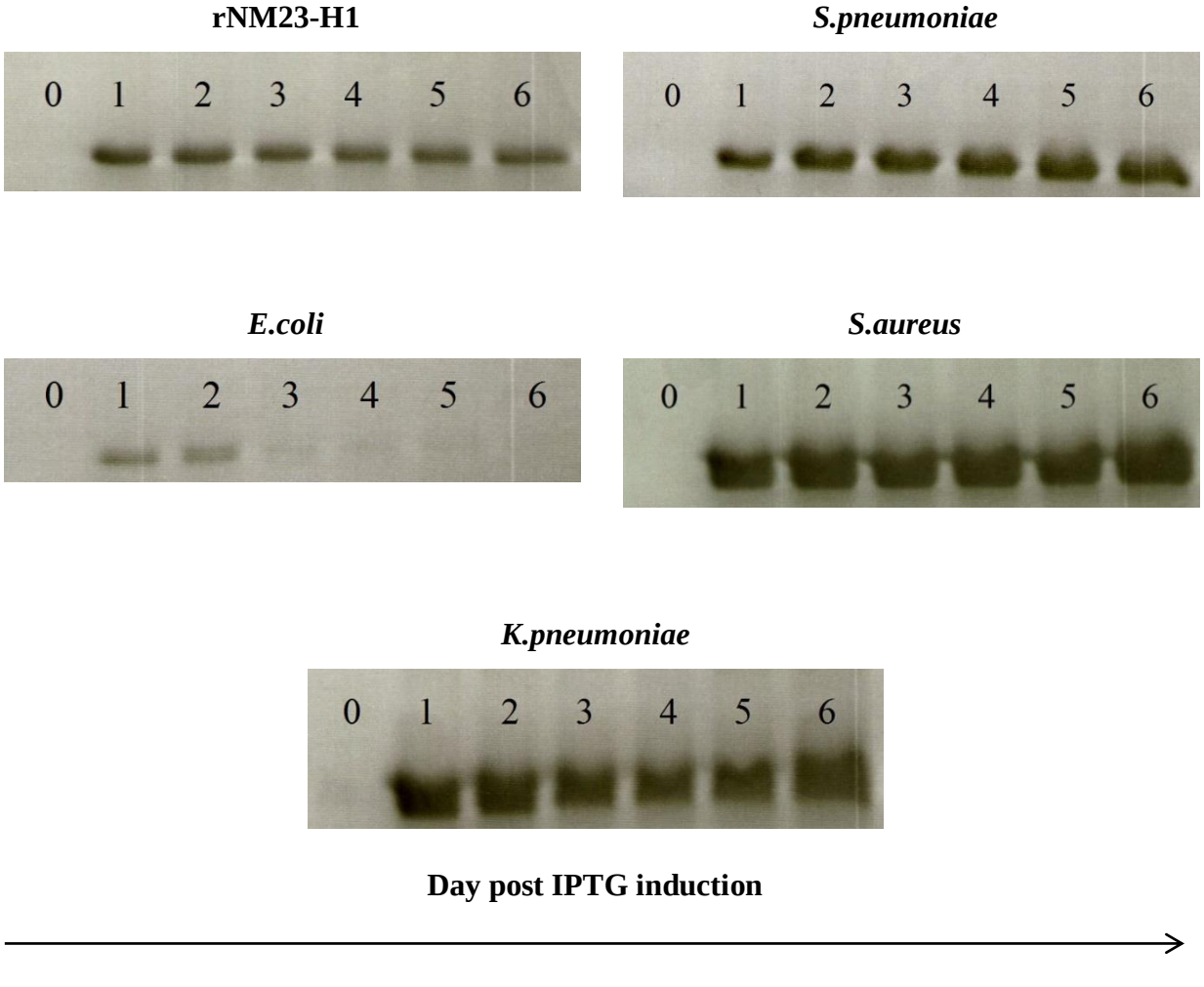
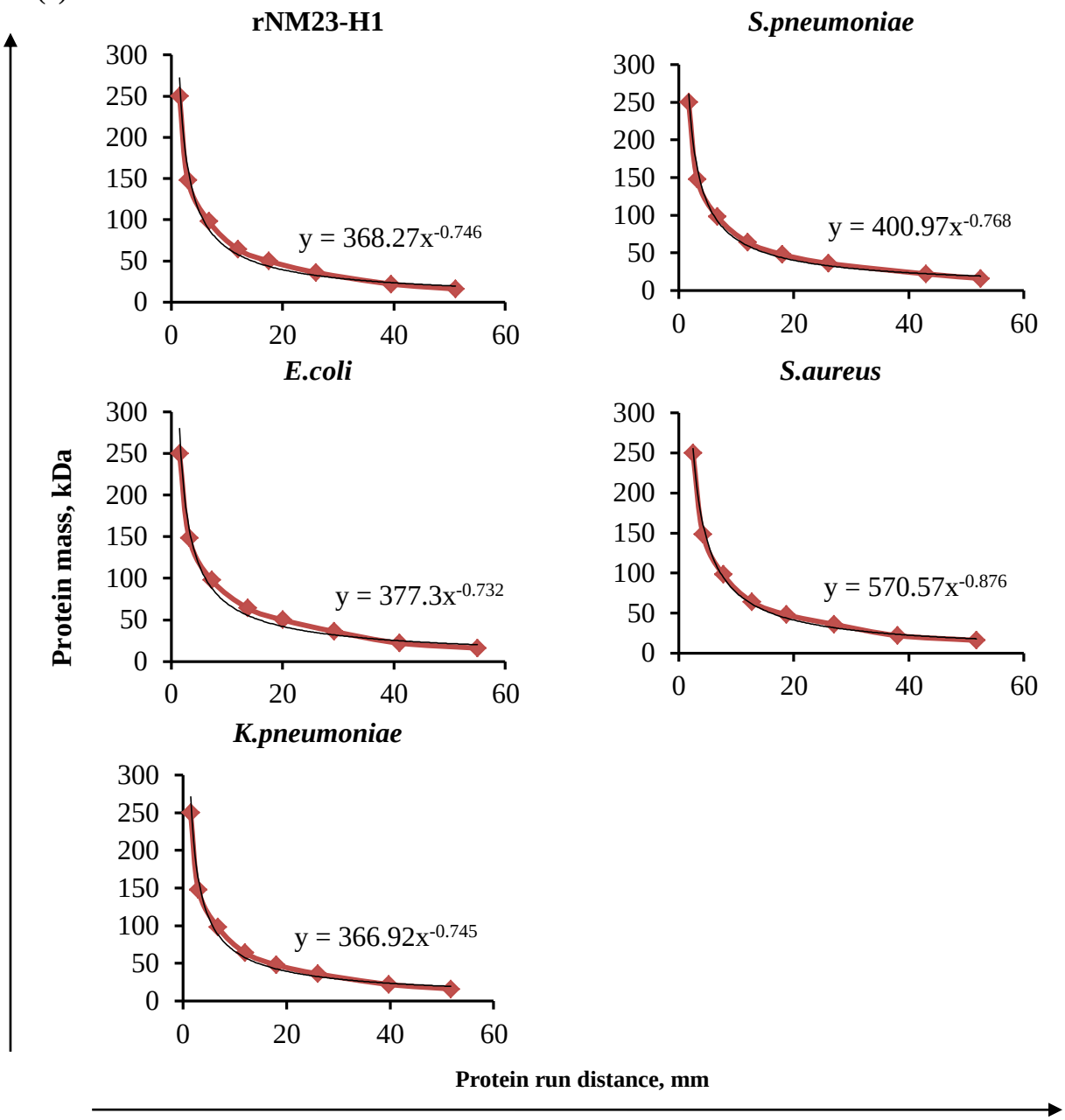


Figure 3.5 Time course experiment: BL21 (DE3) bacteria secrete rNDK proteins. Coomassie blue staining of SDS-polyacrylamide gels. Overnight cultures were diluted 1:50 and rNDK protein expression was induced by IPTG. Samples were collected before ITPG induction (zero time point) and every 24 hours for six days. Prepared samples of supernatants were loaded on SDS-PAGE gels.

(a)



(b)

Protein	Predicted mass, kDa	Calculated mass, kDa
rNM23-H1	19.78	22.26
E.coli rNDK	18.69	20.92
K.pneumoniae rNDK	18.72	20.05
S.pneumoniae rNDK	18.01	19.21
S.aureus rNDK	19.26	19.94

Figure 3.6 Time course experiment: BL21 (DE3) bacteria secrete rNDK proteins. Verification of secreted rNDK proteins. (a) Curves and trend lines of protein ladder mass with respect to their run distance (b) rNDK protein masses calculated from trend line equation and predicted from protein sequences.

3.3 THE OPTIMIZATION OF CD34^{+ve} CELL PURIFICATION

CB samples were sorted for mononuclear cells (pre CD34^{+ve} sorted cells). Further mononuclear cells were sorted for CD34^{+ve} cells as described in materials and methods. Pre and post CD34^{+ve} sorted samples were stained with CD34 FITC (CB2-CB11) or CD34 APC (CB12) antibodies as described in materials and methods and analysed by flow cytometry.

An example of flow cytometry data for pre and post CD34^{+ve} sorted cells is shown for CB7 sample (Figure 3.7). CB7 sample was double sorted for CD34^{+ve} cells. Flow cytometry data for each sample: CD34^{+ve} pre-sorted, CD34^{+ve} after 1st and 2nd sort are shown as two plots. The first one compares the forward and side scatter of all collected events (Figure 3.7, Column 1). An increase in forward scatter indicates an increase in event size. Whereas an increase in side scatter indicates an increase in granulation of cell cytoplasm, thus indicates more mature cells. Gate R1 enclose all viable cells, while dead cell and debris are found in the bottom left corner. Microbeads form a line in the top left corner. For the pre CD34^{+ve} cell sample was collected 100,000 R1 gated events, for the CD34^{+ve} sorted samples 10,000 R1 gated events. Analysing data of CD34^{+ve} once and twice sorted cells viable cells are situated on the bottom of the plot (Figure 3.7, Column 1). Viable cells of pre CD34^{+ve} sorted sample are more distribute towards the increase in side scatter. The second plot represents all viable cells (R1 gated events) and it compares the side scatter and CD34 FITC stained cells (Figure 3.7, Column 2). Gate R2 enclose CD34^{+ve} cells. In pre CD34^{+ve} sorted sample amount of CD34^{+ve} cells is negligible, it increase with each CD34^{+ve} sort (Figure 3.7, Column 2).

Amount of CD34^{+ve} cells in pre and post CD34^{+ve} cell sort samples was calculated from the ratio between microbeads and R2 gated events (Figure 3.7, Column 2). CD34^{+ve} cells are found as 0.5 – 1.7% of mononuclear cells (Figure 3.8a). Percentage of CD34^{+ve} cell for CB2 and CB3 after single sort is 42.1% and 35.1%, respectively. To increase the purity of CD34^{+ve} cells temperature (6°C - 12°C) of MACS buffer was monitored and CD34^{+ve} cell purity

increased to 47.1 - 65.6% (CB4-CB7). Although, monitoring temperature of MACS buffer an increase in purity of $CD34^{+ve}$ cells was observed, the purity of $CD34^{+ve}$ cells is unsatisfactory. Therefore, double sort was performed and an increase in $CD34^{+ve}$ purity 75.8 – 88.9% is observed (Figure 3.8 b).

To visualize diversity of sorted cells, Jenner Giemsa staining was performed (Figure 3.9). Jenner Giemsa stains nuclei blue violet, erythrocytes – weak red, platelets - light pale pink, lymphocyte cytoplasm - sky blue, monocyte cytoplasm – greyish blue, granules in granulocytes – dark violet and stem cells ($CD34^{+ve}$) – violet. Mononuclear layer consist of cells with round nucleus, such as lymphocytes, monocytes or macrophages. In $CD34^{-ve}$ cell sample lymphocytes (dark blue nuclei with a narrow rim of cytoplasm), monocytes (greyish blue) and stem cells (perfectly round and violet) are distinguished. In $CD34^{+ve}$ cell sample most of cells are stem cells ($CD34^{+ve}$) (Figure 3.9).

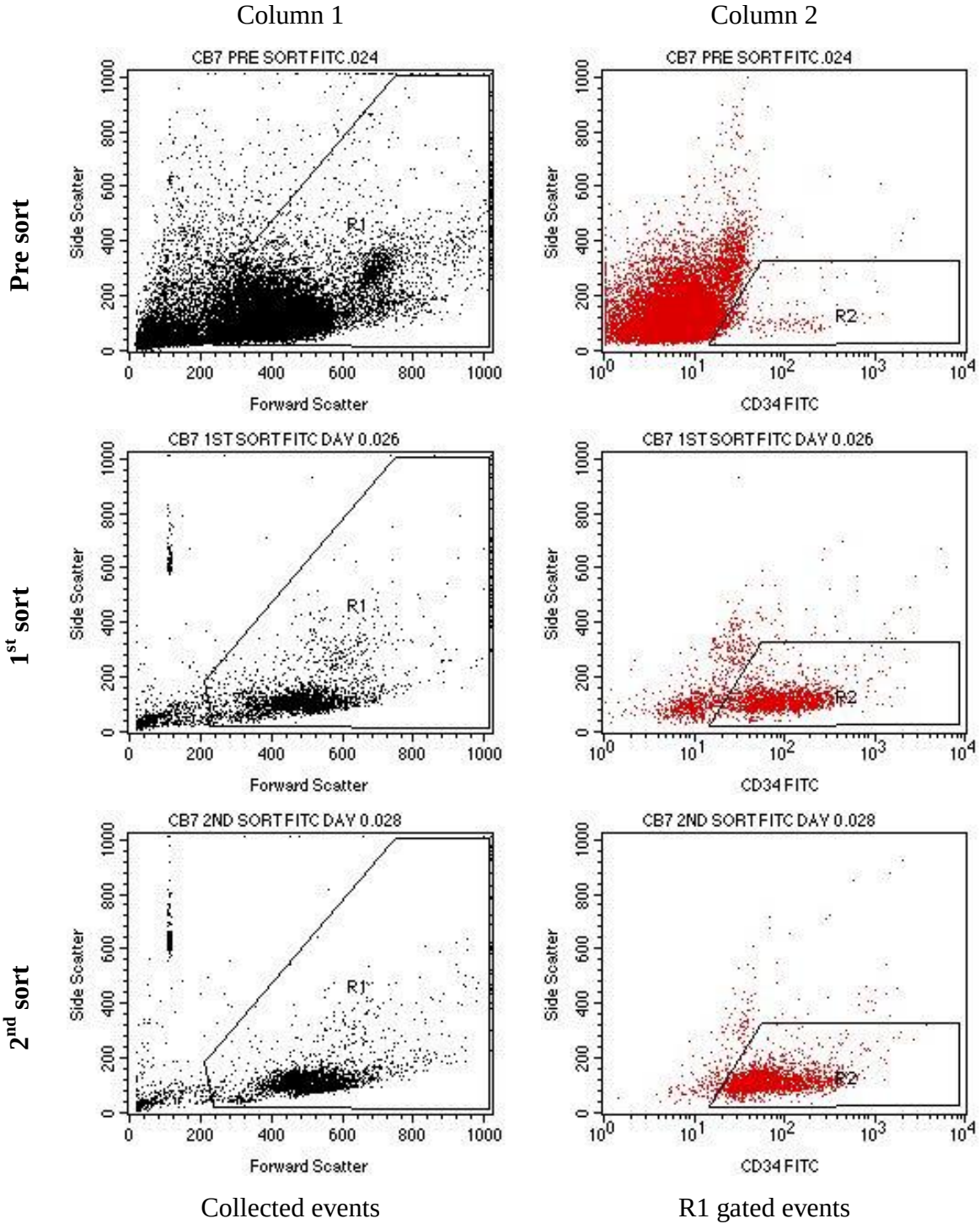
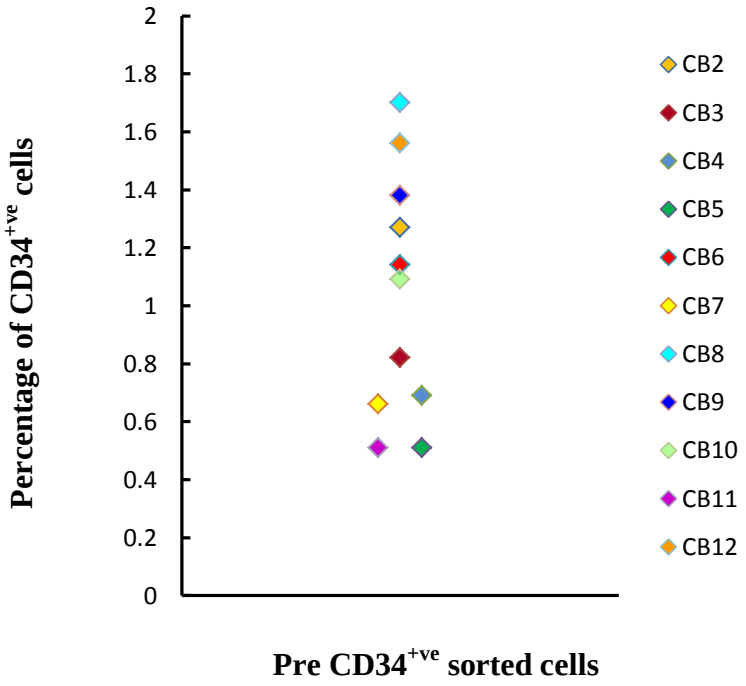


Figure 3.7 CD34^{+ve} cell purification. Flow cytometry data. Pre CD34^{+ve} sorted cells and CD34^{+ve} cells after 1st and 2nd sort (CB7). Total collected events are observed in column 1 plots. Gate R1 enclose viable cells and gate R2 CD34^{+} cells.

(a)



(b)

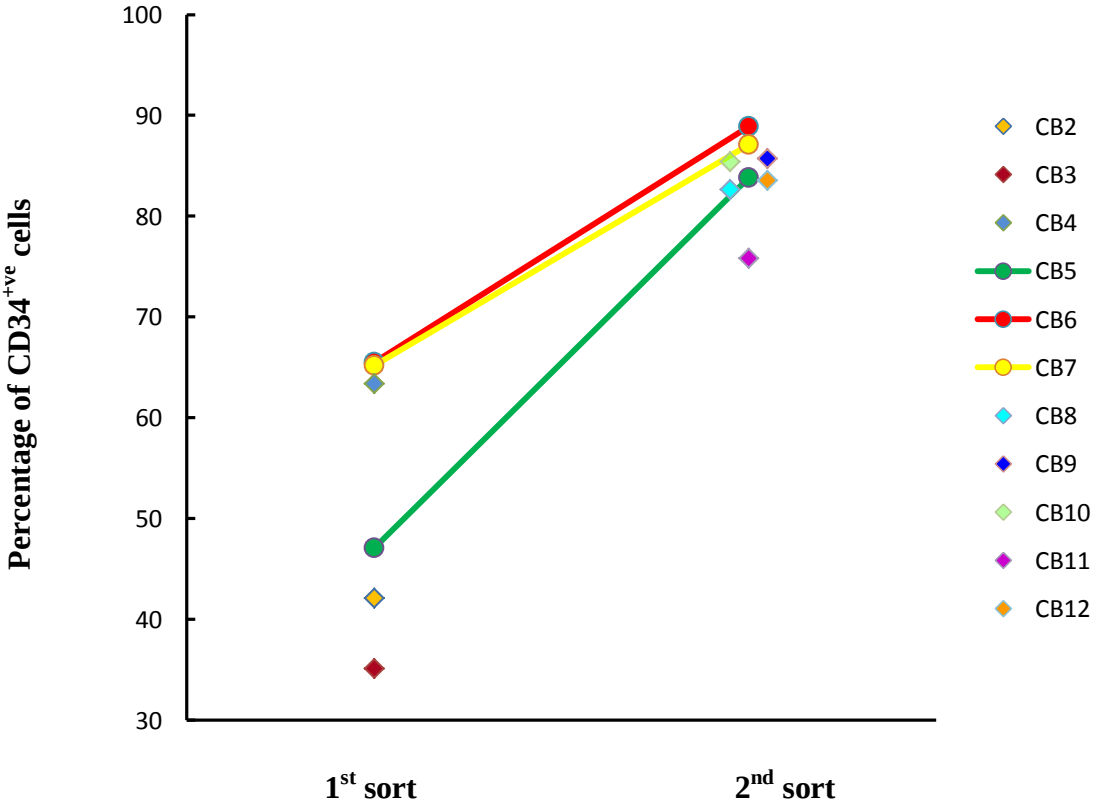
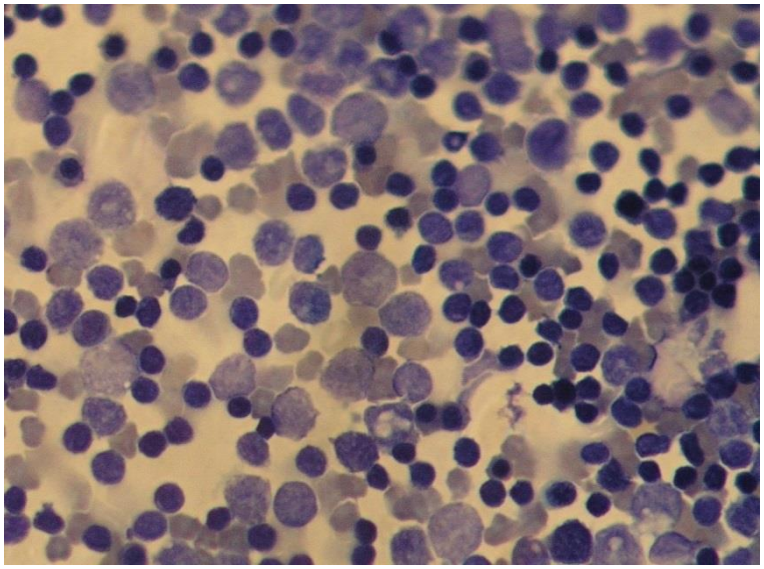


Figure 3.8 CD34^{+ve} cell purification. Percentage of CD34^{+ve} cells. (a) Percentage of CD34^{+ve}cells in pre CD34^{+} sort cells (b) Percentage of CD34^{+ve}cells in CD34^{+ve} cells after 1st and 2nd sort.

CD34^{-ve} cells



CD34^{+ve} cells

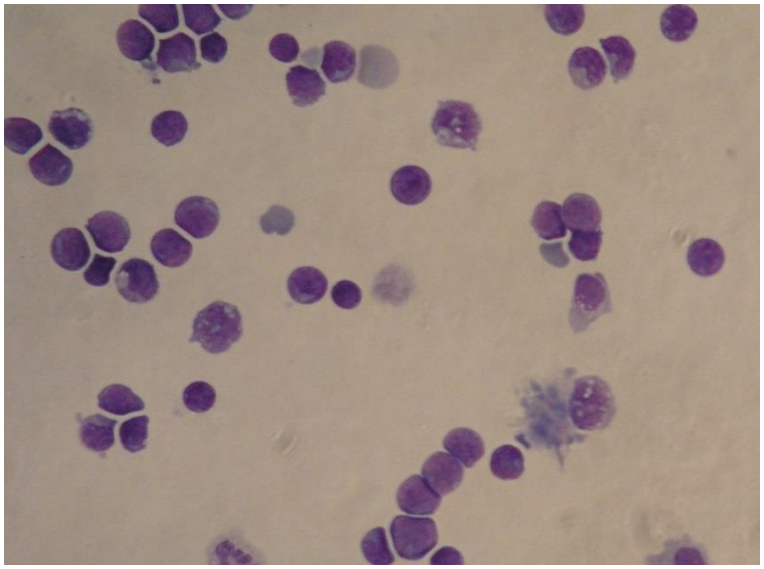


Figure 3.9 CD34^{+ve} cell purification. Post sorted CD34^{-ve} and CD34^{+ve} cell Jenner Giemsa staining (CB2). Pictures are taken at ×400 magnification.

3.4 VIABILITY OF CD34^{-ve} CELLS TREATED WITH rNDK PROTEINS

After CD34 sorting CD34^{-ve} cells were diluted as 0.5×10^6 cells per ml in RPMI 1640 p/s ITS+ and seeded as 500 µl in a 48 well plate. CD34^{-ve} cells were treated with 1µg/ml (CB8-CB10) or 2µg/ml (CB11 and CB12) rNM23-H1 or *E.coli, K.Pneumoniae, S.pneumonia, S.aureus* rNDK protein (rNDK-EC, rNDK-KP, rNDK-SP and rNDK-SA) or EB. CB12 samples were supplemented with 1.25 µg/ml PMB. Treated CD34^{-ve} cells were incubated at 37°C for 7 day. Treated cells were stained with CD34 APC and CD11b PE* antibodies as described in materials and methods and analysed by flow cytometry.

An example of flow cytometry data for treated CD34^{-ve} cell cultures incubated for 7 days is shown for CB11 sample as two plots for each rNDK or EB treatment (Figure 3.10). The first plot of side scatter against forward scatter shows all collected events (Figure 3.10, Column 1). The gate R1 encloses viable cells. The second plot of CD34 APC against CD11b PE* is R1 gated, thus represents all viable cells (Figure 3.10, Column 2). Gate R2 encloses CD11b^{+ve} cells, whilst gate R3 encloses CD34^{+ve} cells. Plots in column 2 shows that CD34^{+ve} and CD11b^{+ve} cells are present in the treated CD34^{-ve} cell cultures. The amount of CD34^{+ve} cells is 0.1-4.0%, although it was expected to be close to 0. Amount of CD11b^{+ve} cells is higher 5.3–40.1% (Figure 3.11). The amount of CD11b^{+ve} cells increases about two times in the sample supplemented with PMB (CB12) (Figure 3.11). Otherwise there is no consistency in the levels of CD11b^{+ve} and CD34^{+ve} cells in treated CD34^{-ve} cells cultures (Figure 3.11)

Figure 3.12 represents viability of treated CD34^{-ve} cells. Total concentration of viable CD34^{-ve} cells decreases from initial concentration 10^6 cells per ml (day 0) to $1.83\pm0.53\times10^5$ (EB), $1.21\pm0.32\times10^5$ (rNM23-H1), $1.10\pm0.26\times10^5$ (rNDK-EC), $1.74\pm0.38\times10^5$ (rNDK-KP), $0.99\pm0.20\times10^5$ (rNDK-SP), $1.59\pm0.45\times10^5$ (rNDK-SA) cells per ml. Total cell concentration

is higher in EB treated samples (negative control) than in rNM23-H1 treated samples (positive control) for all CB samples indicating that positive control failed to work.

$CD34^{-ve}$ cell differentiation was observed for all treated CB samples. $CD34^{-ve}$ cell culture photographs were taken on day 7 (Figure 3.13). $CD34^{-ve}$ cells treated with EB are less differentiated, whilst all rNDK proteins promoted macrophage and fibroblast formation. Cluster formation in $CD34^{-ve}$ cell cultures treated with rNDK was observed only for CB1 and CB11. Cells treated with EB did not form cluster or they were present in smaller amount. Presence and absence of clusters in $CD34^{-ve}$ cell cultures treated with rNM23-1 and EB, respectively, is shown in photographs of Jenner Giemsa stained $CD34^{-ve}$ cells (Figure 3.14).

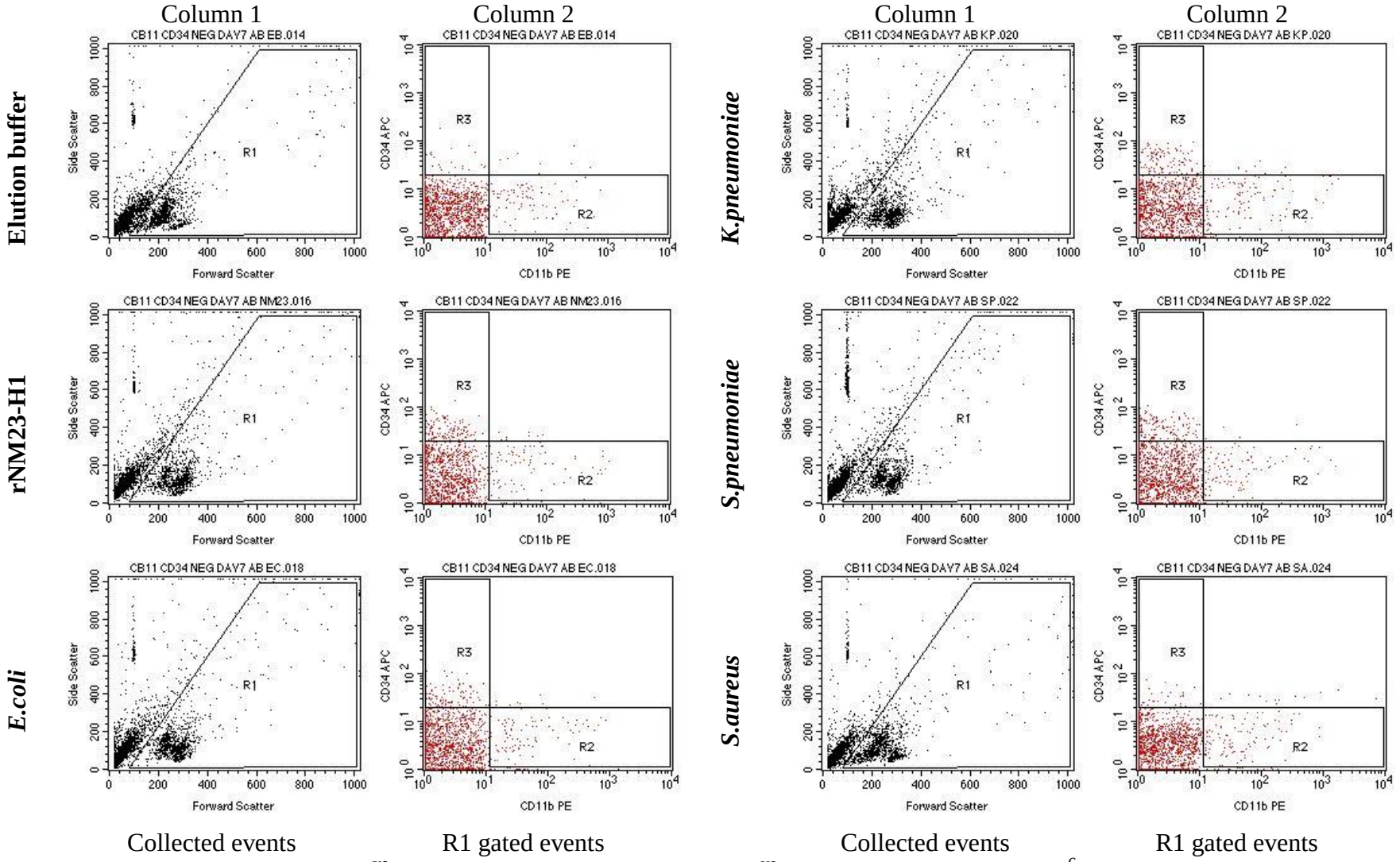
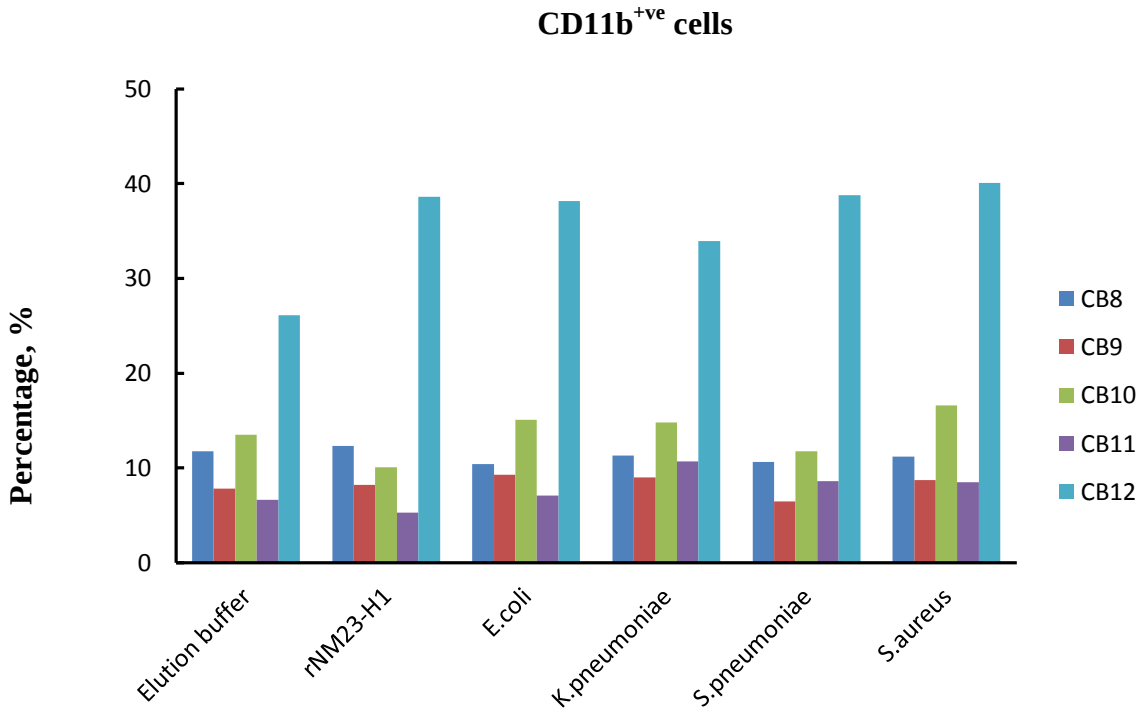


Figure 3.10 Differentiation of CD34^{-ve} cells. Flow cytometry data. CD34^{-ve} cells were seeded as 0.5×10^6 cells per well in a 48 well plate as 500 µl. Seeded cells were treated with 2 µg/ml rNDK protein or EB and incubated at 37ºC for 7 days (CB11). Gate R1 enclose viable cells, gate R2 CD11b^{+ve} cells and gate R3 CD34^{+ve} cells.

(a)



(b)

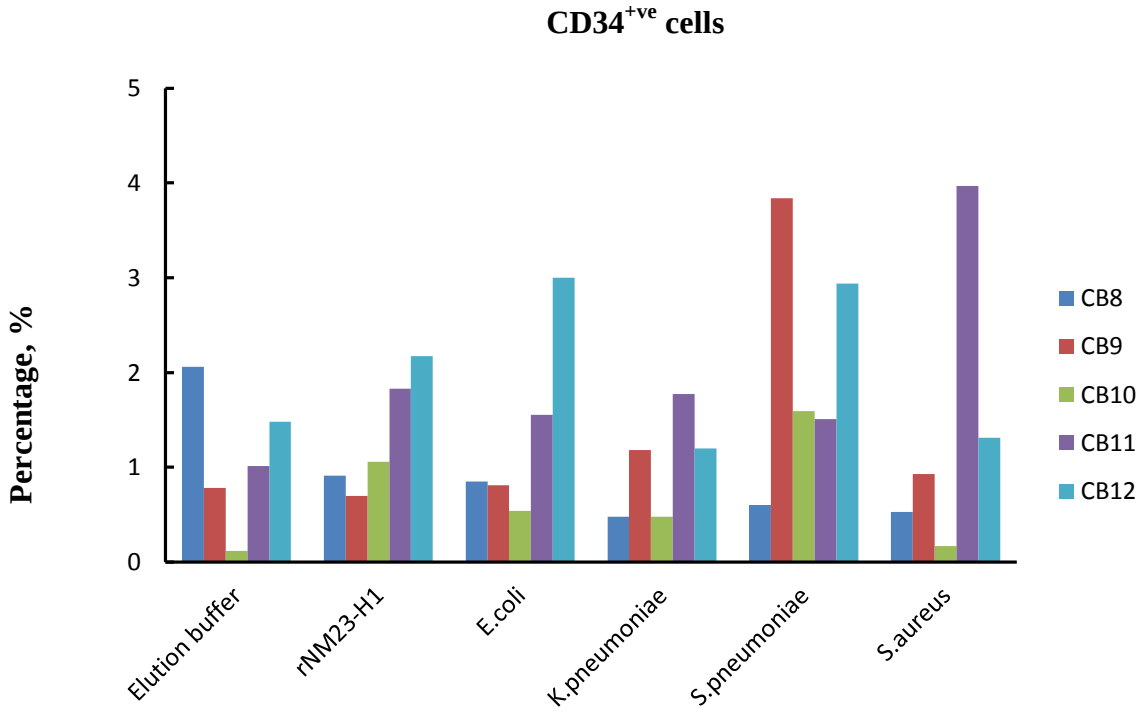


Figure 3.11 CD34^{-ve} cell differentiation. Percentage of CD34^{+ve} and CD11b^{+ve} in treated CD34^{-ve} cell cultures. CD34^{-ve} cells were seeded as 0.5×10^6 cells per well in a 48 well plate as 500 µl. Seeded cells were treated with 1µg/ml (CB8 - CB10) or 2 µg/ml (CB11 and CB12) rNDK protein or EB and incubated at 37ºC for 7 days (CB11). CB12 was treated with 1.25 µg/ml PMB.

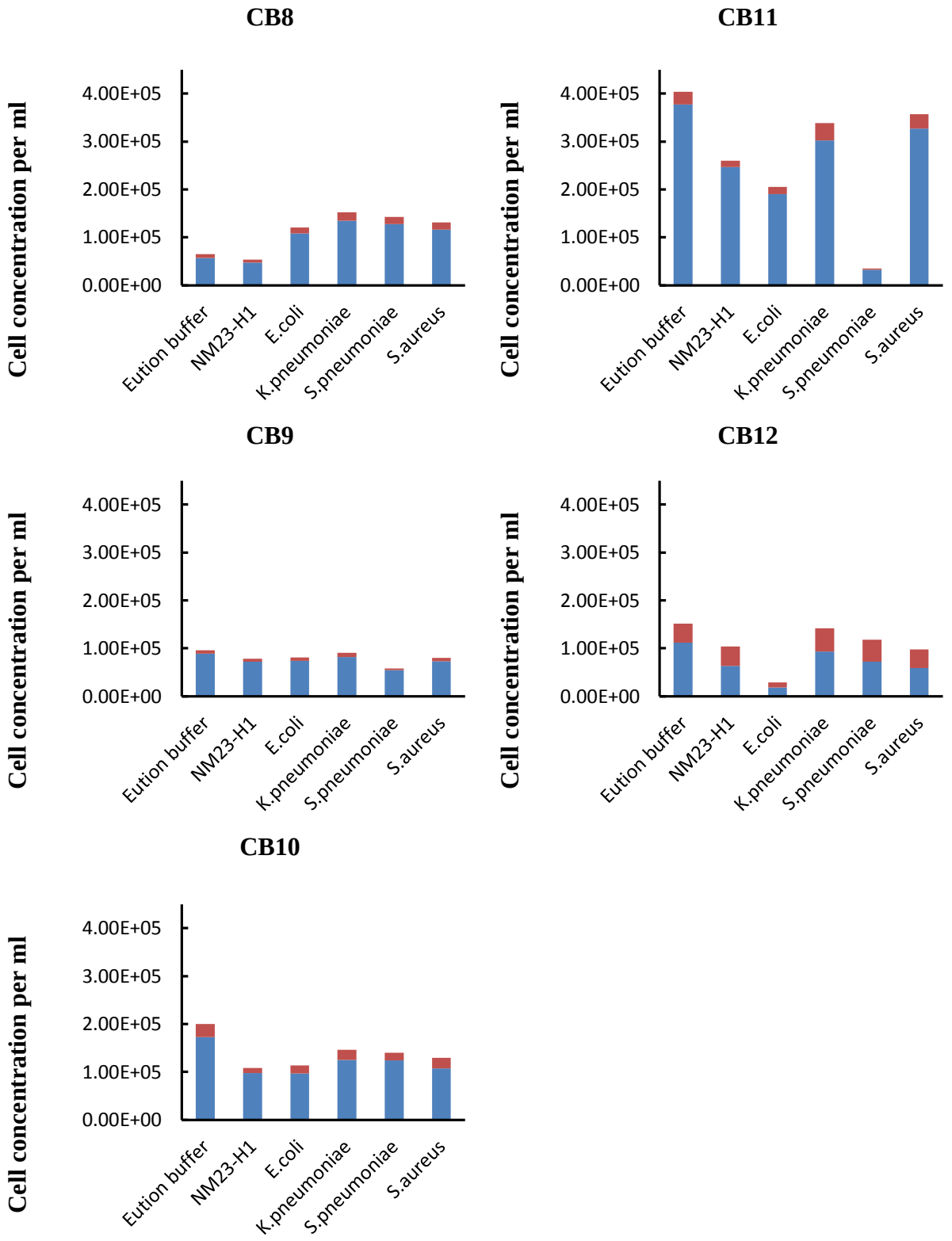
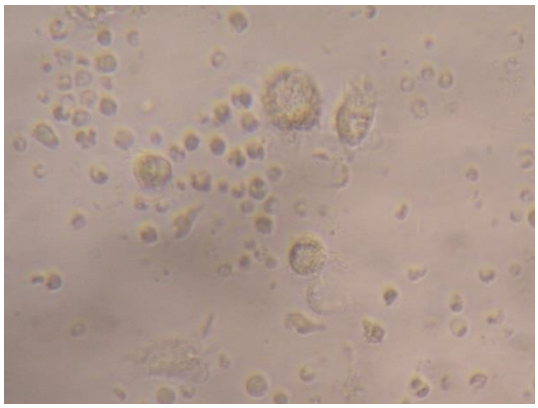
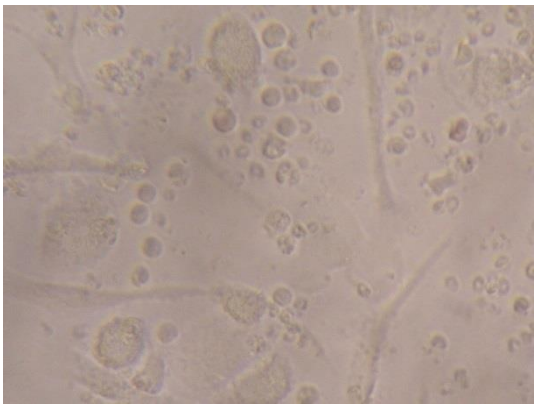


Figure 3.12 CD34^{-ve} cell viability (CD11b^{+ve} cell concentration per ml is represented in red, total cell concentration is sum of blue and red). CD34^{-ve} cells were seeded as 0.5×10^{6} cells per well in a 48 well plate as 500 µl. Seeded cells were treated with 1µg/ml (CB8 - CB10) or 2 µg/ml (CB11 and CB12) rNDK protein or EB and incubated at 37ºC for 7 days (CB11). CB12 was treated with 1.25 µg/ml PMB.

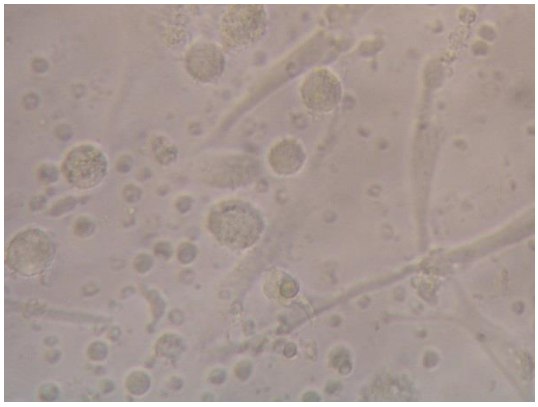
Elution Buffer



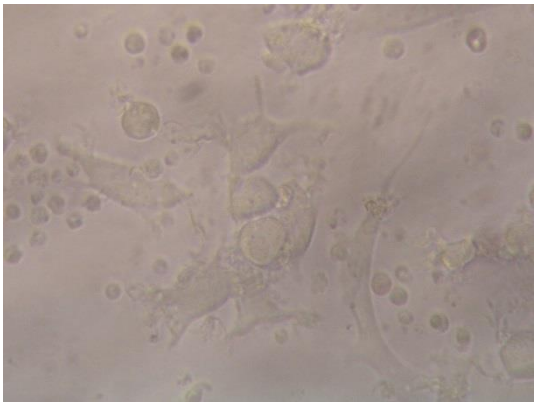
rNM23-H1



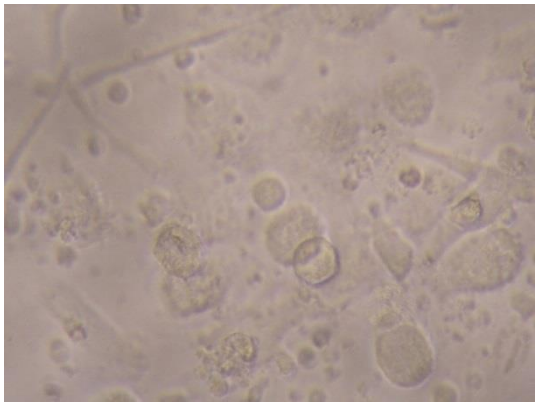
E.coli



K.pneumoniae



S.pneumoniae



S.aureus

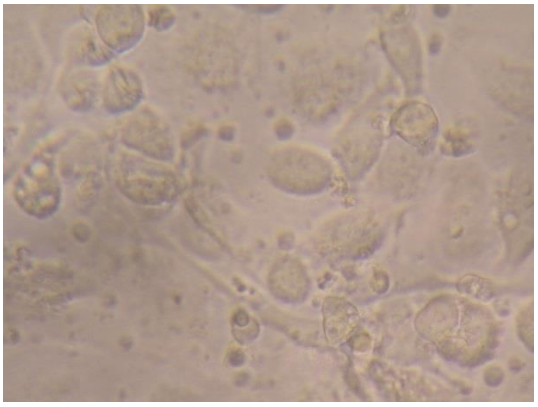
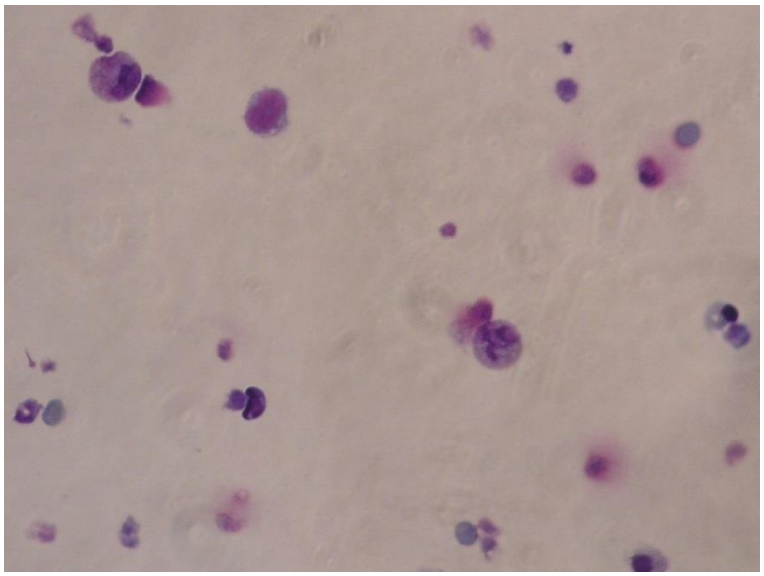


Figure 3.13 CD34^{-ve} cell viability. CD34^{-ve} cells were seeded as 0.5×10^6 cells per well in a 48 well plate as 500 µl. Seeded cells were treated with 2 µg/ml (CB11) rNDK protein or EB and incubated at 37ºC (CB11) for 7 days. Photographs were taken at x400 magnification.

Elution buffer



rNM23-H1

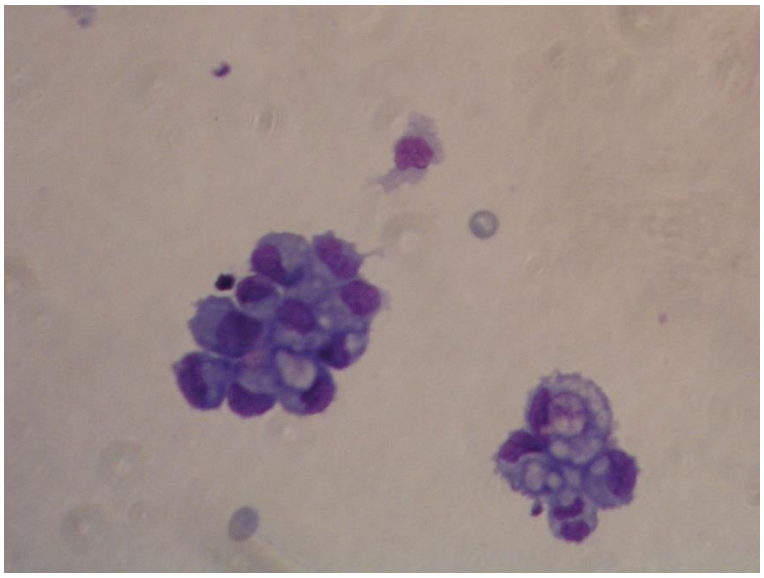


Figure 3.14 CD34^{-ve} cell viability. Photographs of Jenner Giemsa stained CD34^{-ve} cell treated with rNM23-H1 or EB. CD34^{-ve} cells were seeded as 0.5×10^6 cells per well in a 48 well plate as 500 µl. Seeded cells were treated with 2 µg/ml (CB1) rNm23-H1 protein or EB and incubated at 37ºC for 7 days (CB1). Collected cells were cytospined and stained in Jenner Giemsa stain. Photographs were taken at x400 magnification.

3.5 VIABILITY OF CD34^{+ve} CELLS TREATED WITH CM

From CD34^{-ve} cell cultures treated with rNM23-H1 or *E.coli, K.Pneumoniae, S.pneumonia, S.aureus* rNDK protein or EB and incubated for 20 hours CM (CM-NM, CM-EC, CM-KP, CM-SP, CM-SA or CM-EB) was collected. Recombinant NDK proteins were depleted from collected CM as described in materials and methods. Protein depletion was verified by Western blot (Figure 3.15).

CD34^{+ve} sorted cells were left overnight at 4ºC. On the next day CD34^{+ve} cells were counted and diluted as 10^6 cells per ml in CM_{dep}. Resuspended cells were seeded as 150 µl (CB8) and as 100 µl (CB9-CB12) in a 96 well plate and incubated at 37ºC for 7 days. Treated cells were stained with CD34 FITC (CB8-CB11) or CD34 APC (CB12) antibodies as described in materials and methods and analysed by flow cytometry.

An example of flow cytometry data for treated CD34^{+ve} cell cultures incubated for 7 days is shown for CB8 sample as two plots for each CM_{dep}-NDK and CM_{dep}-EB treatment (Figure 3.16). The first plot of side scatter against forward scatter shows all collected events (Figure 3.16, Column 1). Gate R1 encloses viable cells. Amount of viable cells in CD34^{+ve} resuspended in CM_{EB} (CD34^{+ve}/CM_{dep}-EB) is smaller than in CD34^{+ve}/CM_{dep}-NM. The second plot of side scatter against CD34 FITC is R1 gated, thus all events represents all viable cells (Figure 3.16, Column 2). Gate R2 encloses CD34^{+ve} cells. Amount of CD34^{+ve} cells in CD34^{+ve}/CM_{dep}-EB is smaller than in CD34^{+ve}/CM_{dep}-NM (Figure 3.16, Column 2).

Figure 3.17 represents viability of treated CD34^{+ve} cell cultures. CD34^{+ve} cell culture were seeded at total cell concentration 10^6 cells per ml. CD34^{+ve} cell purity for CB8, CB9, CB10, CB11 and CB12 is 82.6%, 85.7%, 85.4%, 75.8% and 83.5%, respectively. Therefore, genuine concentration of CD34^{+ve} cells in seeded samples is 8.26×10^5 (CB8), 8.57×10^5 (CB9), 8.54×10^5(CB10), 7.58×10^5(CB11) and 8.35×10^5(CB12) cell per ml. Total concentration of all,

including $CD34^{+ve}$ cells, from initial concentration 10^6 cells per ml decreases to $1.45\pm0.52\times10^5$ (CM_{dep}-EB), $2.38\pm0.73\times10^5$ (CM_{dep}-NM), $2.46\pm0.87\times10^5$ (CM_{dep}-EC), $2.18\pm1.25\times10^5$ (CM_{dep}-KP), $4.04\pm3.64\times10^5$ (CM_{dep}-SP) and $2.75\pm0.68\times10^5$ cells per ml (CM_{dep}-SA). Whereas concentration of $CD34^{+ve}$ cells only from initial concentration $8.26\pm0.16\times10^5$ cells per ml decreases to $0.59\pm0.12\times10^5$ (CM_{dep}-EB), $1.16\pm0.37\times10^5$ (CM_{dep}-NM), $1.18\pm0.50\times10^5$ (CM_{dep}-EC), $1.33\pm0.78\times10^5$ (CM_{dep}-KP), $2.51\pm1.91\times10^5$ (CM_{dep}-SP), $1.44\pm0.43\times10^5$ (CM_{dep}-SA) Moreover, percentage of $CD34^{+ve}$ cells in treated cultures decreases from 75.8-85.7% to 58.3 - 78.7% (CB8), 45.4 - 50.0% (CB9), 47.3 - 59.0% (CB10), 19.6 - 27.4% (CB11) and 60.1 - 70.7% (CB12). $CD34^{+ve}$ cell viability is not consistent. Comparing viability of $CD34^{+ve}$ cells treated with positive (CM_{dep}-NM) and negative control (CM_{dep}-EB) only in samples CB8, CB11 and CB12 $CD34^{+ve}$ cell concentration is higher in $CD34^{+ve}$/CM_{dep}-NM cells. In CB9 sample concentration of $CD34^{+ve}$ cells treated with CM_{dep}-NM or CM_{dep}-EB are similar, while in sample CB10 $CD34^{+ve}$ cell concentration is higher for CM_{dep}-EB.

Photographs of the treated $CD34^{+ve}$ cell cultures were taken on day 7 (Figure 3.18). In $CD34^{+ve}$/CM_{dep}-EB sample there are fewer cells, whilst in $CD34^{+ve}$/CM_{dep}-NDK are more cells. $CD34^{+ve}$/CM_{dep}-NDK cells are tending to form clusters, whilst in $CD34^{+ve}$/CM_{dep}-EB no cluster or less in comparison to $CD34^{+ve}$/CM_{dep}-NDK is observed. This tendency is observed even when amount of the total cells is very small (CB9 and CB10). Lack and presence of clusters in $CD34^{+ve}$/CM_{dep}-EB and $CD34^{+ve}$/CM_{dep}-NM is observed in Jenner Giemsa stained $CD34^{+ve}$ cell cultures (Figure 3.19).

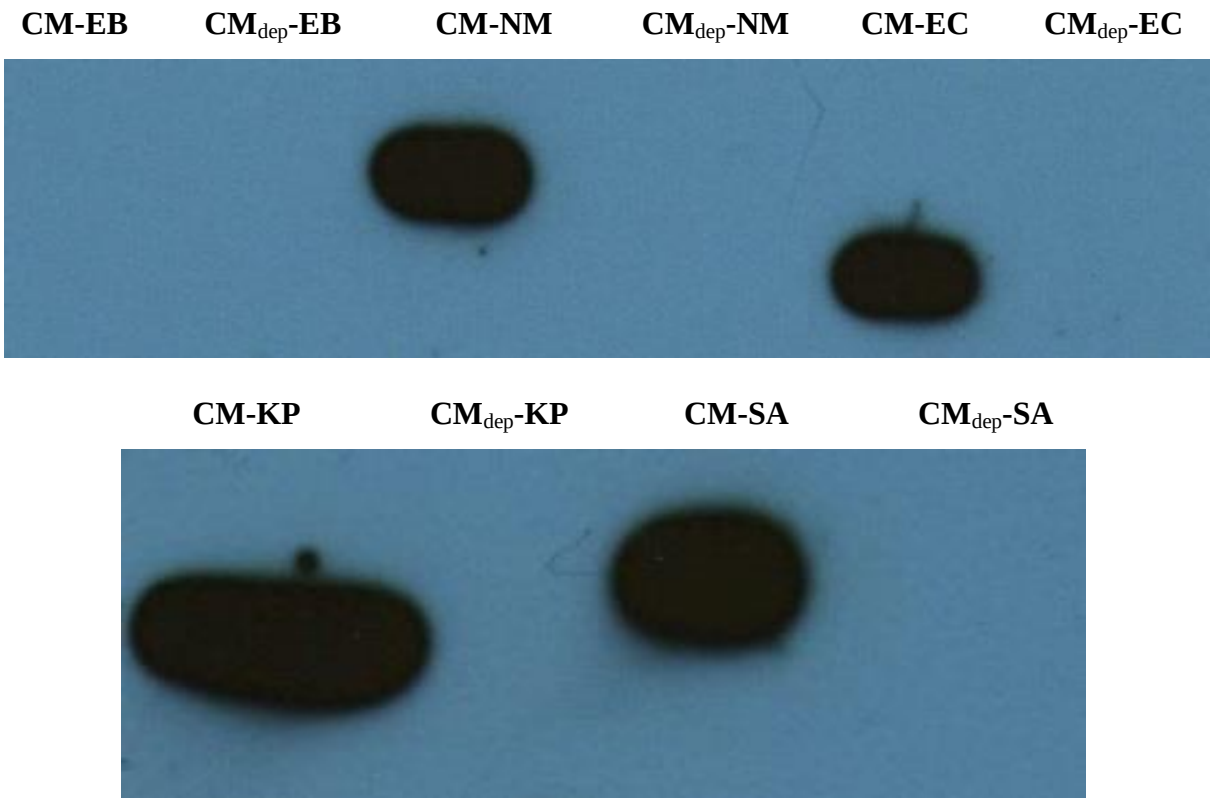


Figure 3.15 Verification of rNDK protein depletion in CM. Western blot analysis (CB12).

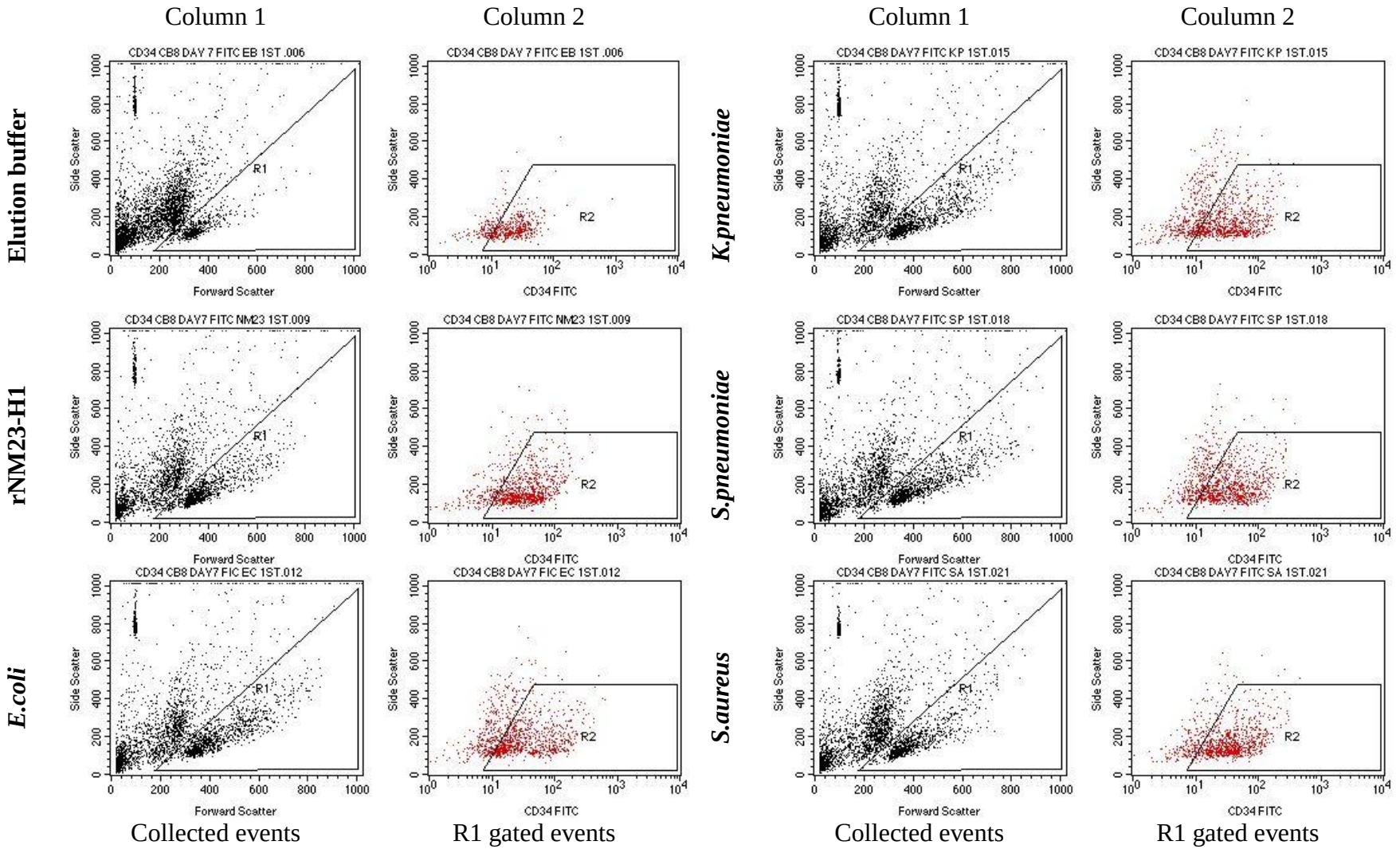


Figure 3.16 CD34^{+ve} cell viability. Flow cytometry data. CD34^{+ve} cells were diluted as 10^6 cells per ml in 150 µl CM_{dep}-NDK or CM_{dep}-EB and seeded in a 96 well plate. Treated CD34^{+ve} cell cultures were incubated at 37ºC for 7 days (CB8). Gate R1 enclose viable cells, gate R2 CD34^{+ve} cells.

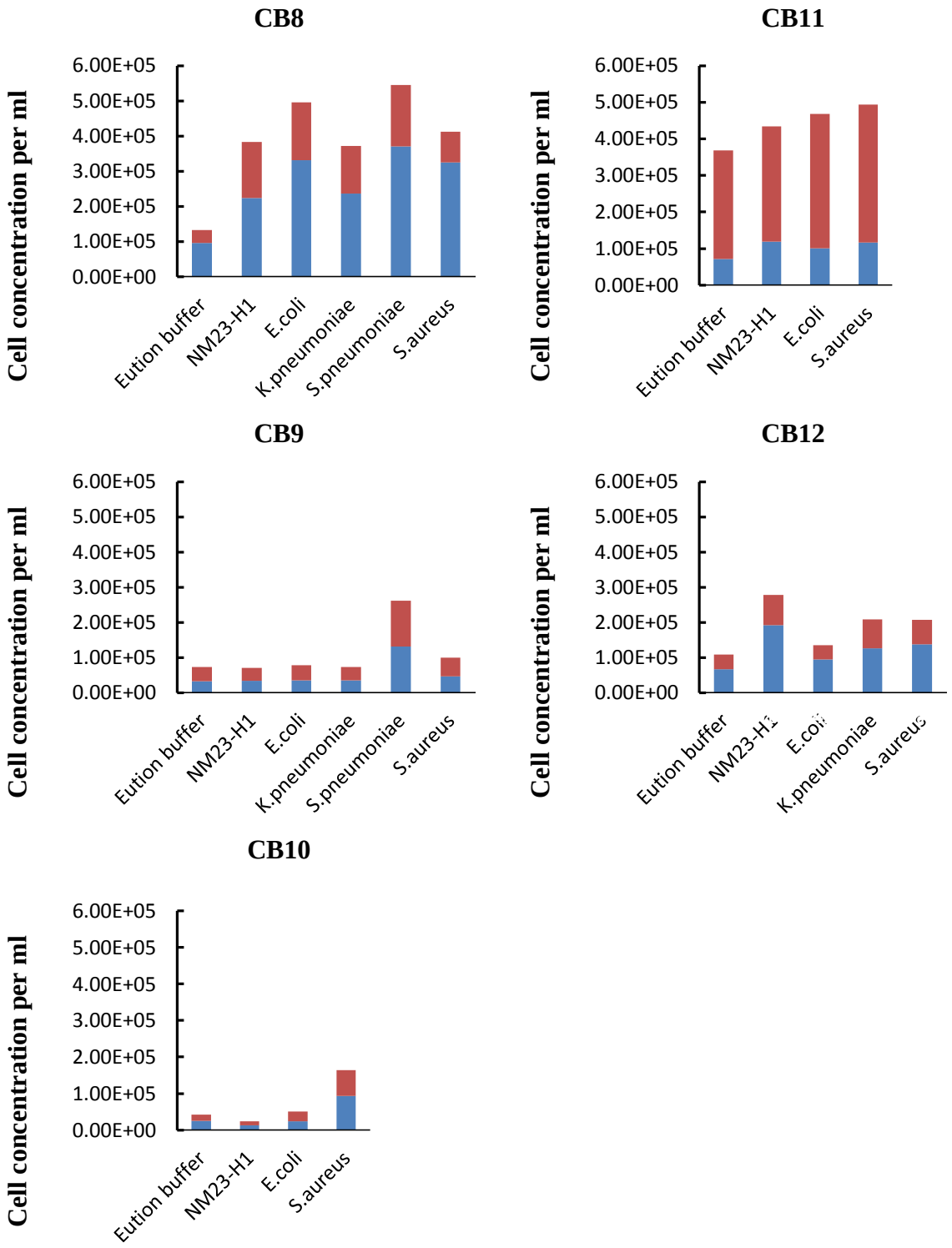
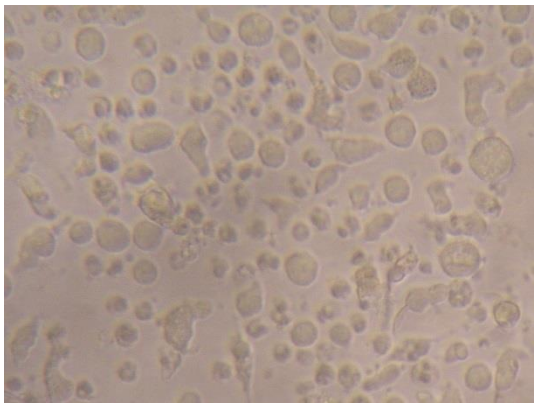
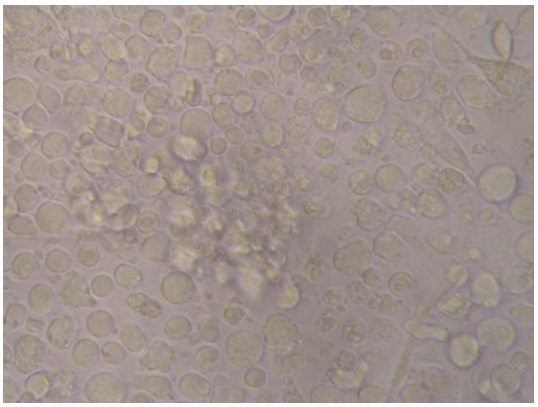


Figure 3.17 $CD34^{+ve}$ cell viability ($CD34^{+ve}$ cell concentration per ml is represented in blue, total cell concentration is in sum of blue and red). $CD34^{+ve}$ cells were diluted as 10^6 cells per ml in 150 µl (CB8) or 100 µl (CB9 – CB12) CM_{dep}-NDK or CM_{dep}-EB and seeded in a 96 well plate. Treated $CD34^{+ve}$ cell cultures were incubated at 37ºC for 7 days.

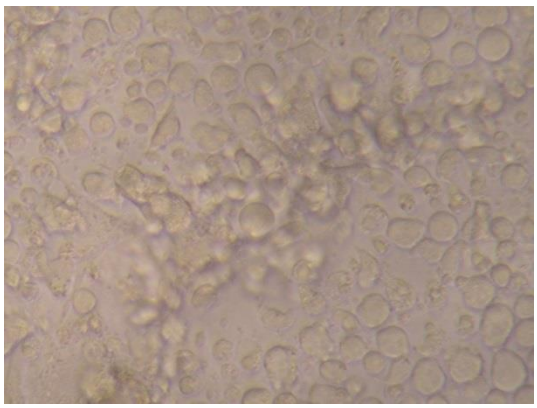
CM_{dep}-EB



CM_{dep}-NM



CM_{dep}-EC



CM_{dep}-SA

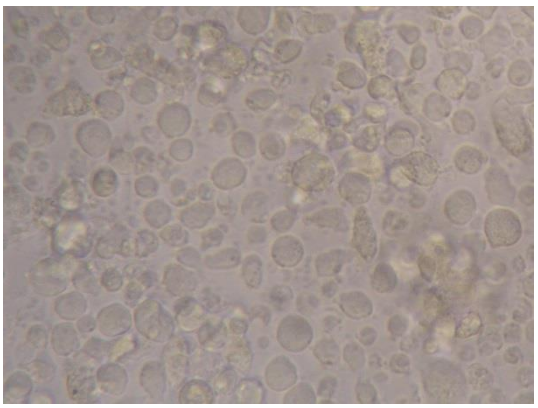
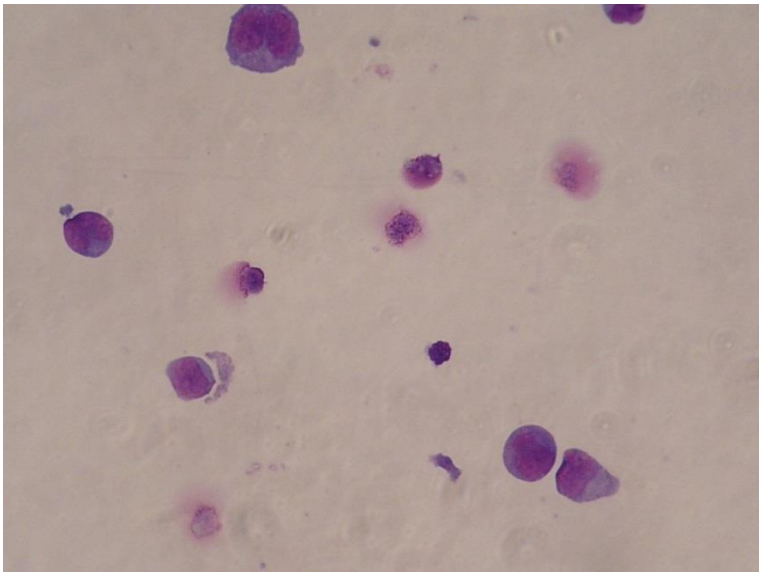


Figure 3.18 $CD34^{+ve}$ cell viability. Photographs of treated $CD34^{+ve}$ cell cultures. $CD34^{+ve}$ cells were diluted as 10^6 cells per 1 ml in 100 µl of CM_{dep}-NDK or CM_{dep}-EB and seeded in a 96 well plate. Treated $CD34^{+ve}$ cell were incubated at 37ºC for 7 days (CB11). Photographs were taken at x400 magnification.

CM_{dep}-EB



CM_{dep}-NM

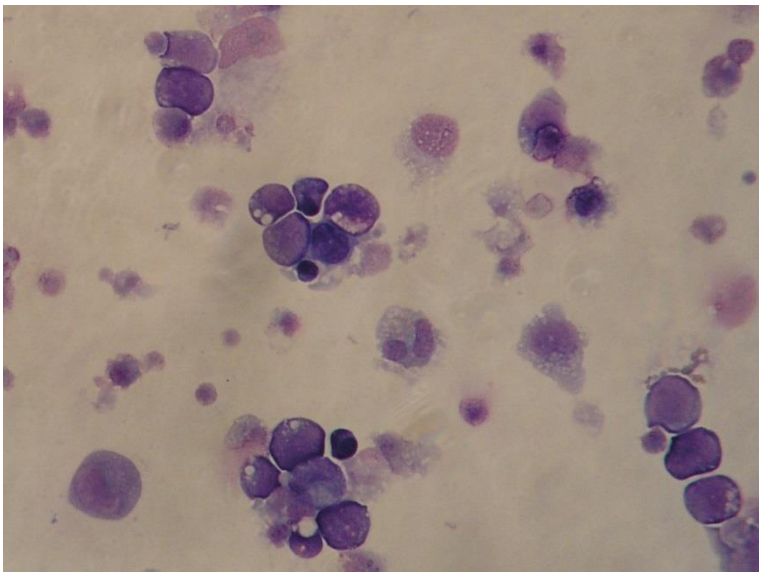


Figure 3.19 Photographs of Jenner Giemsa stained $CD34^{+ve}$ cell treated with CM_{dep}-NM or CM_{dep}-EB (CB1). $CD34^{+ve}$ cells were diluted as 10^6 cells per 1 ml in 200 µl CM_{dep}-NM or CM_{dep}-EB and seeded in a 96 well plate. Treated $CD34^{+ve}$ cell cultures were incubated at 37ºC (CB11) for 7 days. Collected cells were cytospined and stained in Jenner Giemsa stain. Photographs were taken at x400 magnification.

4 DISCUSSION

4.1 BL21 (DE3) BACTERIA SECRETE rNDK PROTEINS

It is known that several pathogenic bacterial taxa such as *P.gingivalis* (Yilmaz et al., 2008), *L.amazonensis* (Kolli et al., 2008), *M. tuberculosis* (Zhao et al., 2007), *P.aeruginosa* (Zaborina et al., 1999), *V.cholera* and *S.typhimurium* (Dar et al., 2011), secrete NDK proteins. The exact mechanism of protein secretion remains hitherto unknown, yet it has been suggested that NDK protein secretion is promoted by the type I mechanism (Kamath et al., 2000). The DXXX (X is predominantly hydrophobic amino acid) motif in the C-terminal region of NDK proteins seems to play a crucial role in NDK protein secretion.

In this project, we could show that *E.coli* BL21 (DE3) secretes rNM23-H1, *E.coli, K. pneumoniae, S.pneumoniae* and *S.aureus* rNDK proteins. All these rNDK proteins possess the sequences DSVE (rNM23-H1, *E.coli* and *K.pneumoniae*), DSLE (*S.aureus*), DSEE (*S.pneumoniae*) or DYTS (rNM23-H1) in the C-terminal end, which might be responsible for their secretion. The secretion mechanism itself was not investigated, as it was not the purpose of this study. An important fact here is that rNDK proteins were secreted from the same *E.coli* BL21 (DE3) strain. Thus, it might be that *K. pneumoniae, S. pneumonia* and *S. aureus* bacteria do secrete NDK proteins expressed by themselves, an observation that should be investigated further.

NM23-H1 and NM23-H2 are secreted by blood cells (Yokoyama et al., 1996, Willems et al., 1998, Okado-Kabe et al., 2002), as they are found in both healthy and in AML patient plasma (Niitsu et al., 2000, Willems et al., 1998, Okado-Kabe et al., 2002). The levels of NM23-H1 and NM23-H2 are increased in leukaemia patients (Niitsu et al., 2000) and

elevated levels of NM23-H1 and NM23-H2 might be just a result of the disease. However, they may as well play a role in the development of leukaemia.

Extracellular expression of NM23-H1 was observed for myeloid and some erythroid cancer cell lines and was not observed for lymphoid cell lines (Okabe-Kado et al., 2002, Willems et al., 1998). Moreover, extracellular levels of NM23-H1 decrease with cell differentiation (Okabe-Kado et al., 2002). Thus, it is possible that by expression of NM23-H1, immature cells affect the functional activity of more mature cells.

4.2 ADDITIONAL rNDK PROTEIN PROMOTES CELL SURVIVAL

Willems et al. (2002) showed that additional NM23-H1, NM23-H2 and NM23-H3 did not affect differentiation of $CD34^{+ve}/CD38^{-ve}$ primary cell culture grown in the presence of IL-1, IL-3, IL-6, kit-ligand and FL cytokines differentiation. Additional NM23 proteins altered differentiation of more mature $CD34^{+ve}/CD38^{+ve}$ cells. It is supported by the findings where additional rNM23-H1 binds better to more mature cells $CD34^{lo}/CD11b^{+ve}$ then to immature $CD34^{+ve}$ (Lilly et al., 2011, Lilly, 2012). An observation of this study is that additional rNM23-H1 protein promoted the survival of both cell populations, although rNM23-H1 predominantly binds to more mature populations. This leads us to suggest that additional rNM23-H1 indirectly promotes immature cell survival. The suggestion was confirmed by an experiment using AML samples. Recombinant NM23-H1 protein indirectly promoted survival of immature $CD34^{+ve}/CD117^{+ve}$.

Bacterial and eukaryotic NDK protein structural similarities (Hama et al., 1991, Chen et al., 2002, Dumas et al., 1992) may mean that NDK proteins are functionally similar. *E.coli, K. pneumoniae, S. pneumoniae* and *Staphylococcus aureus* NDK proteins possess 43 - 53% sequence similarity with NM23-H1. Thus, it might be that bacterial NDK proteins promote AML stem cell survival and proliferation, as shown for rNM23-H1 (Lilly et al., 2011, Lilly, 2012). Bacterial NDK proteins may become a potential target to lessen or eliminate AML. Therefore, it is relevant to test if rNDK proteins are able to promote survival of healthy and cancer blood stem cells.

4.2.1 Direct influence of rNDK proteins on CD34^{-ve} cells

In this part of the project we tested if *E.coli, K. pneumoniae, S. pneumoniae, S.aureus* rNDK proteins promote CD34^{-ve} cell survival. In all experiments, EB was used as a negative control and rNM23-H1 as a positive control. Incubated for 7 days CD34^{-ve} cell were seeded at concentration 10^6 cells per 1 ml after incubating for 7 days the concentration decreased to to $1.83\pm0.53\times10^5$ (EB), $1.21\pm0.32\times10^5$ (rNM23-H1), $1.10\pm0.26\times10^5$ (rNDK-EC), $1.74\pm0.38\times10^5$ (rNDK-KP), $0.99\pm0.20\times10^5$ (rNDK-SP), $1.59\pm0.45\times10^5$ (rNDK-SA) cells per ml. Our positive control treated cells did not survive and did not proliferate, as it was shown by Lilly (2012). Moreover, total cell concentration was higher when CD34^{-ve} cultures were treated with EB in comparison to treatment with rNM23-H1. Unfortunately, due to a failure of our controls, no suggestions can be made about how bacterial rNDK proteins effect CD34^{-ve} cultures. Flow cytometry analysis shows that treated CD34^{-ve} cultures in the presence of EB or rNDK proteins for 7 days have CD34^{+ve} cells (0.1 – 4.0%). The percentage of CD34^{+ve} cells in the CD34^{-ve} fractions was not determined after CD34 sorting. However, the amount of CD34^{+ve} cells in pre CD34^{+ve} sorted cells is about 0.5 – 1.7%. Thus, after CD34 sorting, the percentage of CD34^{+ve} cells in CD34^{-ve} cell fractions should be negligible. The percentage of CD34^{+ve} cells found in CD34^{-ve} cultures is higher than it was determined in pre CD34^{+ve} sort sample, suggesting that CD34^{+ve} cells have proliferated in treated samples. Nevertheless, the CD34^{+ve} cell concentration in treated CD34^{-ve} cultures is just $1.73\pm0.60\times10^3$ (EB), $1.84\pm0.71\times10^3$ (rNM23-H1), $1.27\pm0.43\times10^3$ (rNDK-EC), $2.04\pm0.90\times10^3$ (rNDK-KP), $1.87\pm0.47\times10^3$ (rNDK-SP), $3.43\pm0.24\times10^3$ (rNDK-SA) cells per ml indicating that CD34^{+ve} cells proliferation is putative. It can be assumed that CD34^{+ve} cells left in post sorted CD34^{-ve} sample survived in treated CD34^{-ve} cultures while CD34^{-ve} cells died.

Differentiation of treated $CD34^{-ve}$ cells culture into macrophages and fibroblastic cells for all CB samples was observed. Cell differentiation was confirmed by flow cytometry analysis. We observed that treated $CD34^{-ve}$ cell cultures possess $CD11b^{+ve}$ cells. $CD11b^{+ve}$ cells are found as 5.3 – 40.1% of $CD34^{-ve}$ cell cultures. The amount of $CD11b^{+ve}$ cells increased about two times when $CD34^{-ve}$ cultures were supplemented with PMB (CB12). PMB is used to protect treated cells from bacterial lipopolysaccharides (endotoxins) which could be present in purified rNDK protein fractions. It might be assumed that PMB protects $CD34^{-ve}$ cells from endotoxins, thus allowing cells to differentiate more than it was seen in PMB non-supplemented CB samples. Only one CB sample was supplemented with PMB, so it cannot be suggested if the increased amount of $CD11b^{+ve}$ cells is due to individual properties of CB sample or due to supplemented PMB.

4.2.2 Indirect influence of rNDK protein on $CD34^{+ve}$ cells

Total concentration of all, including $CD34^{+ve}$ cells, from initial concentration 10^6 cells per ml decreases to $1.45\pm0.52\times10^5$ (CM_{dep}-EB), $2.38\pm0.73\times10^5$ (CM_{dep}-NM). Whereas concentration of $CD34^{+ve}$ cells only from initial concentration $8.26\pm0.16\times10^5$ cells per ml decreases to $0.59\pm0.12\times10^5$ (CM_{dep}-EB), $1.16\pm0.37\times10^5$ (CM_{dep}-NM).Although average $CD34^{+ve}$ cells only are higher for the $CD34^{+ve}/CM_{dep}$-NM than it is for $CD34^{+ve}/CM_{dep}$-EB cells, for the CB9 sample was observed the opposite. Higher concentrations of $CD34^{+ve}/CM_{dep}$-NM than of $CD34^{+ve}/CM_{dep}$-EB was observed for samples CB8, CB11 and CB12, whereas for CB10 sample $CD34^{+ve}/CM_{dep}$-EB cells survived just slightly better.

$CD34^{+ve}$ cell cultures resuspended in CM_{dep}-NM (positive control) and incubated for 7days do not survive and do not proliferate as it was shown by Lilly et al (2012). Moreover, comparing the negative control with the positive control our results are not consistent. . As mentioned above our control experiments did not work for $CD34^{-ve}$ cell cultures. $CD34^{-ve}$ cell

cultures did not show the expected survival. Therefore, $CD34^{-ve}$ cell cultures may secrete not enough active components (cytokines), which could cause the expected $CD34^{+ve}$ cell survival and even cell proliferation.

Lilly (2012) previously showed that rNM23-H1 indirectly promotes healthy and leukaemia stem cell ($CD34^{+ve}$) survival. In our experiments, both total cell concentration including $CD34^{+ve}$ cells and concentration of $CD34^{+ve}$ cells only decreased, yet also the percentage of $CD34^{+ve}$ cell in treated $CD34^{+ve}$ cultures decreased. It was expected that even if the total cell concentration decreases, the percentage of $CD34^{+ve}$ cells would stay unchanged or even increased. Nevertheless, the percentage of $CD34^{+ve}$ cells in treated $CD34^{+ve}$ cell cultures decreased from 82.6% to 58.3 - 78.7% (CB8), from 85.7% to 45.4 - 50.0% (CB9), from 85.4% to 47.3 - 59.0% (CB10), from 75.8% to 19.6 - 27.4% (CB11) and from 83.5% to 60.1 - 70.7% (CB12). The highest decrease in $CD34^{+ve}$ cell percentage was observed for CB11. Processed CB11 sample had the smallest purity sorted $CD34^{+ve}$ cells in sorted $CD34^{+ve}$ fraction. It might be that contamination >20% of non $CD34^{+ve}$ cells in $CD34^{+ve}$ seeded cells affected high decrease in $CD34^{+ve}$ cells.

$CD34^{+ve}$ cells resuspended in CM_{dep}-NM and incubated for 7 days tended to form clusters. Treated $CD34^{+ve}$/CM_{dep}-EB cells usually did not form clusters or formed less in comparison to treated $CD34^{+ve}$/CM_{dep}-NM. Cluster formation may indicate that cells are dividing, but cells might just be sticking together. The amount of total cells is decreased. If cells are dividing the speed of new cell formation is much smaller than cell death.

4.2.3 Suggestions why our positive control failed to work

The protocol of cord blood storage used for CB samples by Lilly (2012) was altered in our study. This change could affect $CD34^{-ve}$ and/or $CD34^{+ve}$ cell activity. If $CD34^{-ve}$ cell do not possess their genuine activity, they are not able to release the same amount of cytokines or they may release even different ones.

The fail of positive control can be explained by possible rNM23-H1 activity loss. Aggregation of NDK proteins was observed in the purified NDK protein fractions. It might be that the protein aggregation had changed their active form. Thus, even when proteins are diluted in RPMI 1640 p/s ITS+ to the final concentration of 100 µg/ml, they do not form their native structure. Both suggestions require checking.

4.3 FUTHER STUDIES

Firstly, the reason of the positive control fail requires investigation. When the problem is solved, treatments of $CD34^{-ve}$ and $CD34^{+ve}$ cells purified from CB samples can be performed. If the effect of bacterial rNDK proteins on the survival of healthy $CD34^{-ve}$ and $CD34^{+ve}$ cells has been shown, the bacterial rNDK protein effect on leukaemia stem cells can be better understood.

Lilly et al (2011) reported that $CD34^{-ve}/CD117^{-ve}$ cell cultures treated with rNM23-H1 secrete IL-1β and IL-6 cytokines stronger then $CD34^{-ve}/CD117^{-ve}$ cell cultures treated with EB. Thus, $CD34^{+ve}/CD117^{+ve}$ proliferation can be due to cell response on these cytokines. Similar to findings IL-1β promoted $CD34^{+ve}$ cell proliferation (Ezaki et al., 1995) and IL-6 accompanied with GM-CSF or IL-3 promoted proliferation of normal haemopoietic cell from bone marrow and proliferation of leukaemia myeloid cell (Hoang et al., 1988). Therefore, it should be further investigated if bacterial rNDK proteins will promote cytokine secretion from $CD34^{-ve}$ cell cultures treated with bacterial rNDK proteins and indirect $CD34^{+ve}$ cell survival and proliferation.

If rNDK proteins promote indirect $CD34^{+ve}$ cell survival and proliferation through rNDK protein binding to more mature $CD34^{-ve}$ cells, the most intriguing future direction would be to target these rNDK proteins with the aim of preventing their binding to leukaemia $CD34^{-ve}$ cells, thus preventing cytokine release and leukaemia $CD34^{+ve}$ cell survival.

5 REFERENCES

BACKER, J. M., MENDOLA, C. E., KOVESDI, I., FAIRHURST, J. L., O'HARA, B., EDDY, R. L., JR., SHOWS, T. B., MATHEW, S., MURTY, V. V. & CHAGANTI, R. S. (1993) Chromosomal localization and nucleoside diphosphate kinase activity of human metastasis-suppressor genes NM23-1 and NM23-2. *Oncogene,* 8**,** 497-502.

BERG, J. M., TYMOCZKO, J. L. & STRYER, L. (2002) *Biochemistry,* WH Freeman and Company.

BILITOU, A., WATSON, J., GARTNER, A. & OHNUMA, S. (2009) The NM23 family in development. *Mol Cell Biochem,* 329**,** 17-33.

BOURS, M. J., SWENNEN, E. L., DI VIRGILIO, F., CRONSTEIN, B. N. & DAGNELIE, P. C. (2006) Adenosine 5'-triphosphate and adenosine as endogenous signaling molecules in immunity and inflammation. *Pharmacol Ther,* 112**,** 358-404.

BUCANEVE, G., CASTAGNOLA, E., VISCOLI, C., LEIBOVICI, L. & MENICHETTI, F. (2007) Quinolone prophylaxis for bacterial infections in afebrile high risk neutropenic patients. *Ejc Supplements,* 5**,** 5-12.

BURNETT, A. K., GOLDSTONE, A. H., STEVENS R. M. F., HANN, I. M., REES, J. K. H., GRAY, R. G. & WHEATLEY K. (1998) Randomised comparison of addition of autologous bone-marrow transplantation to intensive chemotherapy for acute myeloid leukaemia in first remission: results of MRC AML 10 trial. *Lancet,* 351, 700-708.

BUSKE, C., FEURING-BUSKE, M., ABRAMOVICH, C., SPIEKERMANN, K., EAVES, C. J., COULOMBEL, L., SAUVAGEAU, G., HOGGE, D. E. & HUMPHRIES, R. K. (2002) Deregulated expression of HOXB4 enhances the primitive growth activity of human hematopoietic cells. *Blood,* 100**,** 862-868.

CHAKRABARTY, A. M. (1998) Nucleoside diphosphate kinase: role in bacterial growth, virulence, cell signalling and polysaccharide synthesis. *Mol Microbiol,* 28**,** 875-882.

CHANG, C. L., ZHU, X. X., THORAVAL, D. H., UNGAR, D., RAWWAS, J., HORA, N., STRAHLER, J. R., HANASH, S. M. & RADANY, E. (1994) Nm23-H1 mutation in neuroblastoma. *Nature,* 370**,** 335-336.

CHEN, Y., GALLOIS-MONTBRUN, S., SCHNEIDER, B., VERON, M., MORERA, S., DEVILLE-BONNE, D. & JANIN, J. (2003) Nucleotide binding to nucleoside diphosphate kinases: X-ray structure of human NDPK-A in complex with ADP and comparison to protein kinases. *J Mol Biol,* 332**,** 915-926.

CHEN, Y., MORERA, S., MOCAN, J., LASCU, I. & JANIN, J. (2002) X-ray structure of Mycobacterium tuberculosis nucleoside diphosphate kinase. *Proteins,* 47**,** 556-557.

COULOMBEL, L. (2004) Identification of hematopoietic stem/progenitor cells: strength and drawbacks of functional assays. *Oncogene,* 23**,** 7210-7222.

DAR, H. H., PRASAD, D., VARSHNEY, G. C. & CHAKRABORTI, P. K. (2011) Secretory nucleoside diphosphate kinases from both intra- and extracellular pathogenic bacteria are functionally indistinguishable. *Microbiology,* 157**,** 3024-3035.

DICK, J. E. (2008) Stem cell concepts renew cancer research. *Blood,* 112**,** 4793-4807.

DUMAS, C., LASCU, I., MORERA, S., GLASER, P., FOURME, R., WALLET, V., LACOMBE, M. L., VERON, M. & JANIN, J. (1992) X-ray structure of nucleoside diphosphate kinase. *Embo J,* 11**,** 3203-3208.

FRITSCHE, T. R., HORN, M., SEYEDIRASHTI, S., GAUTOM, R. K., SCHLEIFER, K.-H. & WAGNER, M. (1999) In Situ Detection of Novel Bacterial Endosymbionts of

Acanthamoeba spp. Phylogenetically Related to Members of the Order *Rickettsiales*. *Appl Environ Microbiol*, 65, 206–212.

GOMEZ, L., GARAU, J., ESTRADA, C., MARQUEZ, M., DALMAU, D., XERCAVINS M., MARTI, J. M. & ESTANY, C. (2003) Ciprofloxacin prophylaxis in patients with acute leukemia and granulocytopenia in an area with a high prevalence of ciprofloxacin-resistant *Escherichia coli. Cancer,* 97, 419-424.

GRIGORIADIS, A. E., KENNEDY, M., BOZEC, A., BRUNTON, F., STENBECK, G., PARK, I.H., WAGNER, E.F. & KELLER, G.M. (2010) Directed differentiation of hematopoietic precursors and functional osteoclasts from human ES and iPS cells. *Blood*, 14, 2769-2776.

JANEWAY, C. A. Jr., TRVERS, P., WALPORT, M. & SHLOMCHIK, M. J. (2001) *Immunobiology: The immune system in health and disease*, Fifth ed., New York, Garland Science.

HELMS, M., VASTRUP, P., GERNER-SMIDT, P. & MOLBAK, K. (2002) Excess mortality associated with antimicrobial drug-resistant *Salmonella Typhimurium*. *Emerg Infect Dis,* 8, 490-495.

HEMERRICH, S., YARDEN, Y. & PECHT, I. (1992) A chromoglycate binding protein from rat mast cells of a leukemia line is a nucleoside diphosphate kinase. *Biochemistry,* 31, 4574-4579.

HEPPNER, G. H. & MILLER, B. E (1983) Tumor heterogeneity biological implications and therapeutic consequences. *Cancer Metast Rev,* 2, 5-23.

HOFFBRAND, A. V., MOSS, P. A. H. & PETTIT, J. E. (2008) *Essential Haematology,* Oxford, Blackwell Publishing Ltd.

HOFFMAN, R., BENZ, E. J., SHATTIL, S. J., FURIE, B., SILBERSTEIN, L. E., MCGLAVE, P. & HESLOP, H. (2009) *Haematology: Basic Principles and Practice,* Philadelphia, Churchill Livingstone Elsevier.

HUANG, J. Y., CHANG, T., CHANG, C. Y. & CHEN, C. J. (2005) Crystal structure of nucleoside diphosphate kinase required for coleoptile elongation in rice (Oryza sativa L.). *J Struct Biol,* 150**,** 309-318.

KAMATH, S., CHEN, M. L. & CHAKRABARTY, A. M. (2000) Secretion of nucleoside diphosphate kinase by mucoid Pseudomonas aeruginosa 8821: involvement of a carboxy-terminal motif in secretion. *J Bacteriol*, 182**,** 3826-3831.

KANG, C.-I., KIM, S.-H., KIM, H.-B., PARK, S.-W., CHOE, Y.-J., OH, M.-D., KIM, E.-C. & CHOE K.-W. (2003) *Pseudomonas aeruginosa* bacteremia: Risk factors for mortality and influence of delayed receipt of effective antimicrobial therapy on clinical outcome. *Clin Infect Dis*, 37, 745-751.

KERN, W. V., KLOSE, K., JELLEN-RITTER, A. S., OETHINGER, M., BOHNERT, J., KERN, P., REUTER S., VON BAUM, H. & MARRE, R. (2005) Fluoroquinolone resistance of *Escherichia coli* at a cancer center: epidemiologic evolution and effects of discontinuing prophylactic fluoroquinolone use in neutropenic patients with leukemia. *Eur J Clin Microbiol*, 24, 111-118.KOLLI, B. K., KOSTAL, J., ZABORINA, O., CHAKRABARTY, A. M. & CHANG, K. P. (2008) *Leishmania*-released nucleoside diphosphate kinase prevents ATP-mediated cytolysis of macrophages. *Mol Biochem Parasitol*, 158, 163-175.

LAU, H. Y., CLEGG, S. & MOORE, T. A. (2007) Identification of Klebsiella pneumoniae genes uniquely expressed in a strain virulent using a murine model of bacterial pneumonia. *Microb Pathog,* 42**,** 148-55.

LEUNG, S.-M. & HIGHTOWER, L. E. A 16-kDa protein functions as a new regulatory protein for Hsc70 molecular chaperone and is identified as a member of the Nm23/nucleoside diphosphate kinase family. *J Biol Chem,* 272, 2607-2614.

LILLY, A. J. (2012) The significance of extracellular Nm23-H1 protein in acute myeloid leukaemia and its role in regulating haemopoietic stem cell behaviour. *School of Biosciences.* Birmingham, University of Birmingham.

LILLY, A. J., KHANIM, F. L., HAYDEN, R. E., LUONG, Q. T., DRAYSON, M. T. & BUNCE, C. M. (2011) Nm23-h1 indirectly promotes the survival of acute myeloid leukemia blast cells by binding to more mature components of the leukemic clone. *Cancer Res,* 71**,** 1177-1186.

LILLY, A., JOHNSON, W. E. & BUNCE C. M. (2011) The haematopoietic stem cell niche: new insights into the mechanisms regulating haematopoietic stem cell behaviour. *Stem Cells Int,* 1-10.

LISTER, M. F., SHARKEY, J., SAWATZKY, D. A., HODGKISS, J. P., DAVIDSON, D. J., ROSSI, A. G. & FINLAYSON, K. (2007) The role of the purinergic P2X7 receptor in inflammation. *J Inflamm (Lond),* 4**,** 5.

LOWY, F. D. (2000) Is Staphylococcus aureus an intracellular pathogen? *Trends Microbiol,* 8**,** 341-343.

MA, D., XING, Z., LIU, B., PEDIGO, N. G., ZIMMER, S. G., BAI, Z., POSTEL, E. H. & KAETZEL, D. M. (2002) NM23-H1 and NM23-H2 repress transcriptional activities of nuclease-hypersensitive elements in the platelet-derived growth factor-A promoter. *J Biol Chem,* 277**,** 1560-1567.

MADANI, T. A. (2000) Clinical infections and bloodstream isolates associated with fever in patients undergoing chemotherapy for acute myeloid leukemia. *Infection,* 28, 367-374.

MANDAI, M., KONISHI, I., KOSHIYAMA, M., MORI, T., ARAO, S., TASHIRO, H., OKAMURA, H., NOMURA, H., HIAI, H. & FUKUMOTO, M. (1994) Expression of metastasis-related nm23-H1 and nm23-H2 genes in ovarian carcinomas: correlation with clinicopathology, EGFR, c-erbB-2, and c-erbB-3 genes, and sex steroid receptor expression. *Cancer Res,* 54**,** 1825-1830.

MARTINEZ, J. A., PREVOT, S., NORDLINGER, B., NGUYEN, T. M., LACARRIERE, Y., MUNIER, A., LASCU, I., VAILLANT, J. C., CAPEAU, J. & LACOMBE, M. L. (1995) Overexpression of nm23-H1 and nm23-H2 genes in colorectal carcinomas and loss of nm23-H1 expression in advanced tumour stages. *Gut,* 37**,** 712-720.

MARTNER, A., DAHLGREN, C., PATON, J. C. & WOLD, A. E. (2008) Pneumolysin released during Streptococcus pneumoniae autolysis is a potent activator of intracellular oxygen radical production in neutrophils. *Infect Immun,* 76**,** 4079-4087.

MARTY, F., YEH, W. W., WENNERSTEN, C. B., VENKATARAMAN, L., ALBANO, E., ALYEA, E. P., GOLD, H. S., BADEN, L. R. & PILLAI, S. K. (2006) Emergence of a clinical daptomycin-resistant *Staphylococcus aureus* isolate during treatment of methicillin-resistant *Staphylococcus aureus* bacteremia and osteomyelitis. *J Microbiol,* 44, 595-597.

MEHUS, J. G., DELOUKAS, P. & LAMBETH, D. O. (1999) NME6: a new member of the nm23/nucleoside diphosphate kinase gene family located on human chromosome 3p21.3. *Hum Genet,* 104**,** 454-459.

MILEO, A.M., PIOMBINO, E., SEVERINO, A., TRITARELLI, A., PAGGI, M. G. & LOMBARDI, D. (2006) Multiple interference of the human papillomavirus-16 E7 oncoprotein with the functional role of the metastasis suppressor Nm23-H1 protein. *J Bioenerg Biomembr,* 38, 215-225.

MILLER, D. M., THOMAS, S. D., ISLAM, A., MUENCH, D. & SEDORIS, K. (2011) c-Myc and cancer metabolism. *Clin Cancer Res,* 18**,** 5546-5553.

MOREAU, A., HILL, M., THEBAULT, P., DESCHAMPS, J. Y., CHIFOLEAU, E., CHAUVEAU, C., MOULLIER, P., ANEGON, I., ALLIOT-LICHT, B. & CUTURI, M. C. (2009) Tolerogenic dendritic cells actively inhibit T cells through heme oxygenase-1 in rodents and in nonhuman primates. *FASEB J*, 23, 3070-3077.MUNOZ-DORADO, J., ALMAULA, N., INOUYE, S. & INOUYE, M. (1993) Autophosphorylation of nucleoside diphosphate kinase from Myxococcus xanthus. *J Bacteriol,* 175**,** 1176-1181.

MURAKAMI, M., KAUL, R., KUMAR, P. & ROBERTSON, E. S. (2009) Nucleoside diphosphate kinase/Nm23 and Epstein-Barr virus. *Mol Cell Biochem,* 329, 131-139.

NAZARETH, H., GENAGON, S. A. & RUSSO, T. A. (2007) Extraintestinal pathogenic Escherichia coli survives within neutrophils. *Infect Immun,* 75**,** 2776-2785.

NIITSU, N., OKABE-KADO, J., NAKAYAMA, M., WAKIMOTO, N., SAKASHITA, A., MASEKI, N., MOTOYOSHI, K., UMEDA, M. & HONMA, Y. (2000) Plasma levels of the differentiation inhibitory factor nm23-H1 protein and their clinical implications in acute myelogenous leukemia. *Blood,* 96**,** 1080-1086.

OKABE-KADO, J., KASUKABE, T. & HONMA, Y. (2002) Expression of cell surface NM23 proteins of human leukemia cell lines of various cellular lineage and differentiation stages. *Leuk Res,* 26**,** 569-576.

OKABE-KADO, J., KASUKABE, T., HONMA, Y., HAYASHI, M., HENZEL, W. J. & HOZUMI, M. (1992) Identity of a differentiation inhibiting factor for mouse myeloid leukemia cells with NM23/nucleoside diphosphate kinase. *Biochem Biophys Res Commun,* 182**,** 987-994.

OKABE-KADO, J., KASUKABE, T., HONMA, Y., HAYASHI, M. & HOZUMI, M. (1988) Purification of a factor inhibiting differentiation from conditioned medium of nondifferentiating mouse myeloid leukemia cells. *J Biol Chem,* 263**,** 10994-10999.

OKABE-KADO, J., KASUKABE, T., HONMA, Y., KOBAYASHI, H., MASEKI, N. & KANEKO, Y. (2009) Extracellular NM23 protein promotes the growth and survival of primary cultured human acute myelogenous leukemia cells. *Cancer Sci,* 100**,** 1885-1894.

OKABE-KADO, J., KASUKABE, T., HOZUMI, M., HONMA, Y., KIMURA, N., BABA, H., URANO, T. & SHIKU, H. (1995) A new function of Nm23/NDP kinase as a differentiation inhibitory factor, which does not require it's kinase activity. *FEBS Lett,* 363**,** 311-315.

OTERO, A. S. (2000) NM23/nucleoside diphosphate kinase and signal transduction. *J Bioenerg Biomembr,* 32**,** 269-275.

PEDRA, J. H., CASSEL, S. L. & SUTTERWALA, F. S. (2009) Sensing pathogens and danger signals by the inflammasome. *Curr Opin Immunol,* 21**,** 10-16.

POSTEL, E. H., BERBERICH, S. J., FLINT, S. J. & FERRONE, C. A. (1993) Human c-myc transcription factor PuF identified as nm23-H2 nucleoside diphosphate kinase, a candidate suppressor of tumor metastasis. *Science,* 261**,** 478-480.

POSTEL, E. H. & FERRONE, C. A. (1994) Nucleoside diphosphate kinase enzyme activity of NM23-H2/PuF is not required for its DNA binding and in vitro transcriptional functions. *J Biol Chem,* 269**,** 8627-8630.

QIN, Z., DAI, L., TOOLE, B., ROBERTSON, E.S. & PARSONS, C. (2011) Regulation of Nm23-H1 and cell invasiveness by Kaposi's sarcoma-associated herpesvirus. *J Virol,* 85, 3596-3606.

ROWE, J.M. (2002) Therapy of secondary leukemia. *Leukemia,* 16, 748-750.

SAHA, A. & ROBERTSON, E. S. (2011) Functional modulation of the metastatic suppressor Nm23-H1 by oncogenic viruses. *FEBS Lett,* 585**,** 3174-3184.

SCHIMPFF, S. C., GREENE, F. A. C. P. W. H., YOUNG, V. M., FORTNER, C. L., JEPSEN, L., CUSACK, R. N. N., BLOCK, J. B. & WIERNIK, P. H. (1975) Infection Prevention in Acute Nonlymphocytic Leukemia Laminar Air Flow Room Reverse Isolation with Oral, Nonabsorbable Antibiotic Prophylaxis. *Ann Intern Med,* 82, 351-358.

SCHLICTMAN, D., KUBO, M., SHANKAR, S. & CHAKRABARTY, A. M. (1995) Regulation of nucleoside diphosphate kinase and secretable virulence factors in Pseudomonas aeruginosa: roles of algR2 and algH. *J Bacteriol,* 177**,** 2469-2474.

SEHL, M. E., SINSHEIMER, J. S., ZHOU, H. & LANGE, K. L. (2009) Differential destruction of stem cells implications for targeted cancer stem cell therapy. *Cancer Res,* 69, 948-949.

SHIZURU, J. A., NEGRIN, R. S. & WEISSMAN, I. L. (2005) Hematopoietic stem and progenitor cells: clinical and preclinical regeneration of the hematolymphoid system. *Annu Rev Med,* 56**,** 509-538.

SPOONER, R. & YILMAZ, O. (2012) Nucleoside-diphosphate-kinase: a pleiotropic effector in microbial colonization under interdisciplinary characterization. *Microbes Infect,* 14**,** 228-237.

TANCREDE, C. H. & ANDREMONT, A. O. (1985) Bacterial Translocation and Gram-Negative Bacteremia in Patients with Hematological Malignancies. *J Infect Diss,* 152**,** 99-103.

WANG, J. C. & DICK, J. E. (2005) Cancer stem cells: lessons from leukemia. *Trends Cell Biol,* 15**,** 494-501.

WANG, X., SHOOK, J., EDINGER, M., WARNER, N. & BUSH-DONOVAN, C. (2012) Multiparametric Immunophenotyping of Human Hematopoietic Stem Cells and Progenitor Cells by Flow Cytometry. *BD Bioscience,* 1-16.

WILLEMS, R., SLEGERS, H., RODRIGUS, I., MOULIJN, A. C., LENJOU, M., NIJS, G., BERNEMAN, Z. N. & VAN BOCKSTAELE, D. R. (2002) Extracellular nucleoside diphosphate kinase NM23/NDPK modulates normal hematopoietic differentiation. *Exp Hematol,* 30**,** 640-648.

WILLEMS, R., VAN BOCKSTAELE, D. R., LARDON, F., LENJOU, M., NIJS, G., SNOECK, H. W., BERNEMAN, Z. N. & SLEGERS, H. (1998) Decrease in nucleoside diphosphate kinase (NDPK/nm23) expression during hematopoietic maturation. *J Biol Chem,* 273**,** 13663-13668.

WILSON, A. & TRUMPP, A. (2006) Bone-marrow haematopoietic-stem-cell niches. *Nat Rev Immunol,* 6, 93-106.

WOGNUM, A.W. & SZILVASSY, S.J. (2013) Hematopoietic stem and progenitor cells. *Stem Cell Technologies, Mini Rev,* 1-8.

YAMAGUCHI, A., URANO, T., GOI, T., TAKEUCHI, K., NIIMOTO, S., NAKAGAWARA, G., FURUKAWA, K. & SHIKU, H. (1994) Expression of human nm23-H1 and nm23-H2 proteins in hepatocellular carcinoma. *Cancer,* 73**,** 2280-2284.

YOKOTA, T., ORITANI, K., BUTZ, S., EWERS, S., VESTWEBER, D. & KANAKURA, Y. (2012) Markers for Hematopoietic Stem Cells: Histories and Recent Achievements, Advances in Hematopoietic Stem Cell Research. IN PELAYO, R. (Ed.) *Advances in Hematopoietic Stem Cell Research.* Rikeja, In Tech.

YOKOYAMA, A., OKABE-KADO, J., SAKASHITA, A., MASEKI, N., KANEKO, Y., HINO, K., TOMOYASU, S., TSURUOKA, N., KASUKABE, T. & HONMA, Y. (1996) Differentiation inhibitory factor nm23 as a new prognostic factor in acute monocytic leukemia. *Blood,* 88**,** 3555-3561.

ZABORINA, O., MISRA, N., KOSTAL, J., KAMATH, S., KAPATRAL, V., EL-IDRISSI, M. E., PRABHAKAR, B. S. & CHAKRABARTY, A. M. (1999) P2Z-Independent and P2Z receptor-mediated macrophage killing by Pseudomonas aeruginosa isolated from cystic fibrosis patients. *Infect Immun,* 67**,** 5231-5242.

ZHAO, Q., LOGOTHETIS, D. E., & SEGUELA, P. Regulation of ATP-gated P2X receptors by phosphoinositides. *Pflugers Arch,* 455, 181-185.

PUBMED PROTEIN DATABASE (no date) *RecName: Full=Nucleoside diphosphate kinase 3; Short=NDK 3; Short=NDP kinase 3; AltName: Full=DR-nm23; AltName: Full=Nucleoside diphosphate kinase C; Short=NDPKC; AltName: Full=nm23-H3.* Available at: http://www.ncbi.nlm.nih.gov/protein/Q13232.2 (Accessed 18th September 2013).

PUBMED PROTEIN DATABASE (no date) *nm23-H7 [Homo sapiens]* Available at: http://www.ncbi.nlm.nih.gov/protein/AAD34622.1 (Accessed 18th September 2013).

PUBMED PROTEIN DATABASE (no date) *NM23-H8 [Homo sapiens]*. Available at: http://www.ncbi.nlm.nih.gov/protein/AAF20909.2 (Accessed 18th September 2013).

PUBMED PROTEIN DATABASE (no date) T*hioredoxin domain-containing protein 6 [Homo sapiens]*. Available at: http://www.ncbi.nlm.nih.gov/protein/NP_835231.1 (Accessed 18th September 2013).

6 APPENDIX

6.1 BACTERIAL GROWTH MEDIA

6.1.1 LB

20 g dissolve in up to 1 liter dH_2O

6.1.2 LB agar

1.2% agar in 1 liter of LB-Broth

6.2 PCR

6.2.1 1×PCR mix (50 µl) contained:

27 µl dH_2O,

5 µl 10×PCR buffer (Bioline) ,

1.5 µl 50mM $MgCl_2$ (Bioline),

1 µl Tag polymerase (Bioline),

1.5 µl 10 mM dNTPs.

6.2.2 The sequence of the used primers (Sigma):

5'TAATACGACTCACTATAGGG3' (T7 forward primer),

5'GGCCCGTCTTCACCACAT3' (NM23-H1 reverse primer),

5'GTAGGGATCCTTATTAACGGGTGCGCGGGCACAC3' (*E.coli* rNDK reverse primer),

'5GTAGGGATCCGATTAGCGAGTGCGGGCAAACC3' (*K.pneumonia* rNDK reverse primer),

5'GTAGGGATCCCTCTTAAAACCAAAGAGCAATTTC3' (*S.pneumoniae* rNDK reverse primer),

5'GTAGGGATCCATTTTATTCATATAACCATGCATC3' (*S.aureus* rNDK reverse primer).

6.3 AGAROSE GEL ELECTROPHORESIS

6.3.1 1% agarose gel

1 g agarose dissolve in 100ml 1×TBE buffer heating the mix. When mix cooled add 3 µl ethidium bromide

6.3.2 1×TBE buffer

10.8 g Tris

5.5 g Boric acid

4 ml 5 M EDTA pH 8.0

Make up to 1 liter H_2O

6.4 SDS-PAGE ELECTROPHORESIS

6.4.1 10×runing buffer

30 g Tris (250 mM)

144 g Glycine (1.92 M)

10 g SDS (1%)

Dissolve in up to 1 liter dH_2O

6.4.2 Resolving gel 12.5% (for one gel)

3.2 ml dH_2O

2.5 ml 1.5 M Tris HCl pH 8.8

4.2 ml ProtoGel (30% (w/v) acrylamide: 0.8% (w/v) Bis-acrylamide stack solution (37.5:1)

100 µl 10% SDS

50 µl 10% Ammonium persulphate

3.3 µl N,N,N',N'-tetramethyl-ethylenediamine (TEMED)

6.4.3 Stacking gel (for one gel)

2.03 ml dH_2O

830 µl 0.5 M Tris HCl pH 6.8

440 µl ProtoGel (30% (w/v) acrylamide: 0.8% (w/v) Bis-acrylamide stack solution (37.5:1)

66 µl 10% SDS

16.7 µl 10% Ammonium persulphate

1.7 µl TEMED

6.4.4 4×GLB

3.7 ml dH_2O

1.25 ml 0.5 M Tris HCl pH 6.8 (62.5 mM)

2.5 ml glycerol (40%)

2.0 ml 10% SDS (2%)

0.5 ml 2-β-mercaptoethanol

Brilliant blue

6.4.5 1.5 M Tris (pH 8.8)

90.8 g Tris

Dissolve in 400 ml, adjust pH to 8.8

Make up to 500 ml in dH_2O

6.4.6 0.5 M Tris (pH 6.8)

30.3 g Tris

Dissolve in 400 ml, adjust pH to 6.8

Make up to 500 ml in dH_2O

6.5 WESTERN BLOT

6.5.1 Transfer buffer

7.575 g TRis (25 mM)

36 g glycine (192 mM)

500 ml methanol (20%)

Make up to 2.5 liters with dH_2O

6.5.2 TBS-T

40 ml 1.0 M tris HCl pH 7.6

16 g NaCl

2 ml Tween20

Make up to 2 liters dH_2O

6.5.3 5% Blotto (one membrane)

2 g dried skimmed milk (5% w/v)

40 ml TBS-T

6.6 COOMASSIE STAINING

6.6.1 Coomassie stain:

50% (v/v) methanol

10% (v/v) acetic acid

0.05% (w/v) Coomassie Brilliant blue

6.6.2 Destain

7% (v/v) acetic acid

5% (v/v) methanol

6.7 OTHERS

6.7.1 Equilibration buffer

50mM sodium phosphate (pH8)

0.3M sodium chloride,

10mM imidazole in H2O

6.7.2 FACS Fix

1% formaldehyde (v/v)

2% FBS (v/v)

in 500 ml PBS

6.7.3 Giemsa buffer

8 mM KH_2PO_4

6 mM Na_2HPO_4

Adjust to pH 7.0.

6.7.4 MACS buffer

2 mM EDTA

1% FBS

in PBS

6.7.5 RIPA buffer

1% v/v NP40 (IGEPAL)

0.5% w/v sodium deoxycholate 0.5g

0.1% w/v 10% SDS 1ml

Make up to 100 ml in dH2O