3 Introduction

The objective of this work was to find peptides and peptoids capable of inhibiting cytokines. The focus lies on a special class of cytokines, the so-called chemokines. Section 3.1 gives a basic overview of chemokine structure and function with special focus on the chemokine CXCL8 (see Section 3.1.1). A main function of CXCL8 is the recruitment of leukocytes into inflamed tissue. To enter into the tissue, leukocytes have to leave the blood vessel across the endothelial border. This process is called 'extravasation' and is explained in Section 3.1.2. The directed movement to the site of inflammation is called 'migration' and the underlying cellular processes are summarized in Section 3.1.3. Chemokines are an interesting target for drug development due to their involvement in a number of diseases, some of which are described in Section 3.1.4. An overview of previously reported chemokine inhibitors is given in Section 3.1.5. Possible strategies for the discovery of novel protein ligands by library screening are shown in Section 3.2 with focus on chemical (see Section 3.2.1) and biological (see Section 3.2.2) libraries. Screening success can be verified by testing affinities of novel discovered compounds in ligand binding assays whose basic principles are explained in Section 3.2.3. Fluorescence anisotropy is a simple yet effective method for studying protein:ligand interactions in solution and is described in detail in Section 3.2.3. Peptidomimetics possess a structure similar to peptides but have several advantages like increased stability and better bioavailability when compared to peptides. Peptoids, the class of peptidomimetics used for the discovery of novel CXCL8 ligands in this work, are characterized in Section 3.3.

3.1 Chemokines

The human immune system is divided into two major sectors: the innate immune system and the adaptive immune system. Innate immunity refers to a non-specific defense system which is capable of attacking pathogens such as bacteria or viruses at the site of inflammation. Neutrophil granulocytes (neutrophils), a type of polymorphonuclear leukocytes (PMN) are the predominant cell type acting in the innate immune system and the most abundant type of white blood cells (leukocytes) in the human body.[1] Their defense mechanisms include phagocytosis, respiratory burst (release of reactive oxygen species, ROS) and the degranulation leading to the release of antimicrobial substances and enzymes from their characteristic granula.[2] The peripheral leukocytes routinely circulate the human blood ready to defend the host against pathogens. Given that the intrusion of a pathogen most likely takes place at mucous membranes and skin (in particular damaged parts thereof) the crucial step of the innate immune defense is the contact of leukocytes with pathogens outside the blood vessel. This can only be achieved by an intricate process involving the directed migration of white blood cells out of the artery into adjacent inflamed tissue (extravasation or diapedesis).[3–6] This directed cell movement is orchestrated by chemotactic cytokines, the so-called chemokines. Apart from leukocytes also various other blood cells, epithelial cells, glial cells and muscle cells secrete chemokines upon stimulation by pathogens or proinflammatory cytokines.[7–10]

Chemokines are small proteins of 8-12 kDa. They posess a highly conserved tertiary structure though their sequence can vary significantly.[11] They display high pI values that lead to a positive overall molecule charge under physiological conditions. Chemokines are therefore preset for binding to negatively charged molecules such as glycosaminoglycans (GAG) on cell-surfaces.[12] They interact with G-protein coupled receptors (GPCRs) on leukocytes or endothelial cells to initiate signalling that ultimately leads to integrin activation, actin polymerization and oxidative burst.

To this day around 50 chemokines and 20 chemokine receptors have been discovered in the human body.[11] They are divided into four subclasses dependent on the number and distribution of the cysteine residues at the N-terminus. With zero, one or three amino-acids in between cysteine residues there are CC-chemokines, e. g. CCL5 (RANTES), CCL2 (MCP-1), CCL11 (eotaxin-1), CXC-chemokines, e. g. CXCL8 (interleukin-8), CXCL12 (SDF1-α) and CX_3C-chemokines, e. g. CX_3CL1 (fractalkine). Two C-chemokines with only one cysteine residue have also been discovered: XCL1 (lymphotactin) and XCL2 (SCM1-β). This 'cysteine-code' is followed by an 'L' for *ligand* when referring to a chemokine, or by an 'R' when referring to a *receptor*. CXC chemokines are further divided into ELR^+CXC and ELR^-CXC chemokines, depending on whether they posses the 'ELR-motif', a sequence of glutamic acid, leucin and arginine at the N-terminus (in front of the first N-terminal cysteine residue); ELR^+CXC chemokines display angiogenic activity while ELR^-CXC chemokines display angiostatic activity.[13] The tertiary structure of all known chemokines is highly conserved and consists of an unstructured N-terminus followed by an antiparallel three-stranded β-sheet and a C-terminal α-helix.[14]

3.1.1 The Chemokine CXCL8

CXCL8 (or interleukin-8, IL-8) is a pro-inflammatory chemokine that was first described in 1987 as a tissue-derived neutrophil-activating protein.[15,16] Earlier names for CXCL8 include neutrophil activating factor (NAF), granulocyte chemotactic peptide (GCP) and monocyte-derived neutrophil chemotactic factor (MDNCF). It has a molecular mass of 8386 g/mol and possesses an overall positive charge at physiological pH values due to its high isoelectric point that has been reported to be 8.3[17] and is calculated to be 9.0 with ExPASy.[18] CXCL8 is the best studied member of the subfamily of angiogenic ELR^+CXC chemokines and possesses strong neutrophil activating and recruitment properties. CXCL8 can be produced by leukocytes like neutrophils, monocytes, T-cells, natural killer cells, as well as by endothelial cells, fibroblasts and epithelial cells.[19] CXCL8 is produced as an inactive 99 amino acid precursor with a 20 amino acid signal sequence.[17,20] 79-, 77-, 72-, 71-, 70- and 69-amino-acid isoforms have been isolated from cultured human blood cells.[16,21] The most active form of CXCL8 *in vivo* is the 72 amino acid isoform.[19]

CXCL8 binds two different GPCRs: CXCR1 and CXCR2[32,33] with an affinity of approximately 2 nM for both receptors.[34] CXCR1 is only activated by CXCL8 and CXCL6, CXCR2 is activated by CXCL8 and multiple other chemokines.[35] Monocytes, eosinophils, neutrophils and basophils all express both CXCR1 and CXCR2 receptors but neutrophils show the highest tendency to migrate upon CXCL8 stimulation.[36] While neutrophils express both receptors in roughly the same amounts, CXCR2 seems to be the predominant receptor in basophils, eosinophils and monocytes.[37] The interactions between CXCL8 and its receptor CXCR1 have been studied extensively. From early on, the ELR motif at the CXCL8 N-terminus was discovered to be essential for receptor activation.[29,30] CXCL8(6-72) without the ELR motif still binds to CXCR1 though with reduced affinity.[29] Early mutational studies with CXCR1/2 in-

Table 3.1: Structure and functional intra- and intermolecular interactions of CXCL8.

Part of CXCL8	Length	Part of Sequence	Interactions
α-helix	17 aa	55-72	Interaction with GAG[22–24]
50S loop	4 aa	52-55	
β_3 strand	5 aa	48-51	Disulfide bond with second cysteine (C9),[25] interaction with CXCR1(1-38)[26,27]
40S loop	3 aa	44-47	Interaction with CXCR1(1-38)[26,27]
β_2 strand	6 aa	38-43	
30S loop	9 aa	29-37	Disulfide bond with first cysteine (C7)[25]
β_1 strand	7 aa	22-28	Dimerisation site[28]
3_{10} helix	3 aa	19-21	Interaction with CXCR1(1-38)[26,27]
N-loop	9 aa	10-18	Interaction with CXCR1(1-38),[26,27] Y13-F17 for CXCR1/CXCR2 specificity[14]
N-terminus	9 aa	1-9	Receptor activation,[29,30] possible interaction with Arg199, Arg203, Asp265 of CXCR1[31]

tersubtype chimeric receptors showed, that the N-terminus of the CXCR1/2 receptors was responsible for chemokine specificity and thus was a likely chemokine:receptor interaction site.[38] NMR studies of the N-terminus of CXCR1, CXCR1(1-40) with ^{15}N-labelled CXCL8(1-72) suggested the N-loop, 3_{10} helix, 40S loop, β_3 strand and α-helix as CXCL8:receptor interaction sites.[26] However, the α-helix as interaction site is disputable since studies with truncated CXCL8(1-51) suggested that the α-helix was not necessary for receptor binding.[29,39] Mutational studies of CXCR1 implicated the involvement of E275, R280[40] as well as R199 and R203[31] in CXCL8:receptor binding. These residues are situated near extracellular domains ECD3 and ECD4 of CXCR1 and discussed as a possible interaction site for the ELR motif.[31] Since two interaction sites on CXCR1, N-terminus and ECD were considered likely, a two-site model was suggested for the binding of CXCL8 to CXCR1. In this model the CXCL8 N-loop first interacts with the CXCR1 N-terminus (site I) and the CXCL8 N-terminus containing the ELR motif then interacts with ECD3 and ECD4 (site II).[41–43] Recent NMR studies of CXCL8 with parts of the receptor in a lipid bilayer supported the CXCL8:CXCR1(1-38) interaction reported earlier by Clubb *et al.*[26] but failed to provide experimental proof of the interaction between ELR motif and CXCR1 extracellular domain regions.[27] However, it was lately reported that ECD4 enhances binding affinity of the CXCR1 N-terminus for CXCL8 when ECD4 and N-terminus of CXCR1/2 were fused onto a soluble protein scaffold that retains the orientation of ECD and N-terminus of the chemokine receptor structure original (chemokine receptor elements on a soluble scaffold, *CROSS*).[44] The CROSS protein uniting CXCR1 N-terminus and ECD4 bound CXCL8 with an affinity of 0.8 ± 0.3 μM while ECD4 without an additionally fused CXCR1 N-terminal part did not bind

CXCL8. Despite the lack of definite proof, the hypothesis of synergistic binding effects between two regions of CXCL8 is still considered valuable.

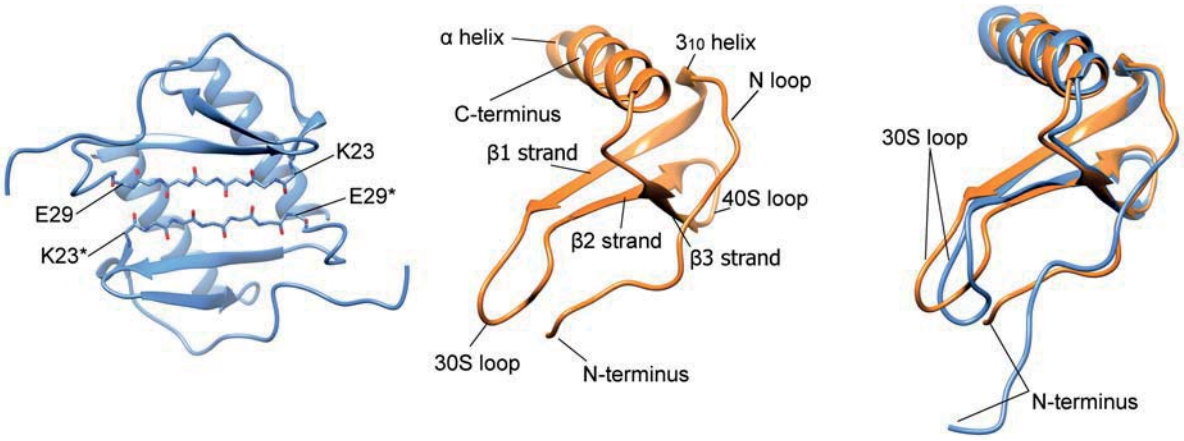

Figure 3.1: Left: CXCL8 dimer structure based on NMR (pdb id: 1IL8)[25] with peptide backbone of the dimerisation interface reaching from residues K23 to E29 (mainly β_1 strand).[28] The sequence misses the first serine S1 of the 72 amino acid isoform and starts with alanine A2. Middle: one chain of the CXCL8 dimer structure based on NMR refined by X-ray data (pdb id: 3IL8),[28] the sequence starts with arginine R6. Right: superimposition of X-ray (orange) and NMR (blue) structure shows significant differences in 30S loop and N-terminus, which are the largely unstructured parts of the chemokine and major CXCL8:CXCR1 interaction sites.

From early on NMR and crystal structures suggested that CXCL8 is capable of forming homodimers.[25,28] The β1 strands of two CXCL8 molecules align to form a six-stranded antiparallel β-sheet stabilizing the dimer (see Figure 3.1).[45] The dimerisation interface comprises mainly the β_1 strand residues K23 to E29.[28] The dimerisation constant of CXCL8 was reported to be between 0.1 μM and 20 μM.[46–48] There is still some debate about the CXCL8 monomer and dimer form and their respective biological activity. The monomeric form of CXCL8 is likely to be the active form that binds CXCR1.[27,47–49] However, recent *in vivo* studies with trapped monomer and trapped dimer variants of CXCL8 show neutrophil recruitment ability for both variants.[50,51] It was hypothesized that CXCL8 dimerisation could lead to negative regulation.[52] A conclusive hypothesis that takes all the findings into account is not yet available.

A monomer/dimer equilibrium can lead to significant complication of experiments, e. g. the evaluation of binding data.[53] To further study the influence of the monomer/dimer equilibrium on CXCL8 function, several different CXCL8 variants have been produced. First, a 'nonassociating' monomer CXCL8-L25NMe with a backbone methylation at L25 was synthesized by solid phase peptide synthesis.[49] NMR studies revealed a loss of helical structure in residues 67-72 (pdb id: 1IKL).[49] It was concluded that the outer end of the C-terminal α-helix of CXCL8 was merely stable when the dimer was formed.[54] Later studies determined the affinity of CXCL8L25NMe for CXCR1 to be $K_d = 0.8 \pm 0.3$ nM, which is equivalent to wild type CXCL8 within the error of the binding assay.[55] Secondly, the non-dissociating dimer CXCL8-R26C with an affinity of $K_d = 55.0 \pm 6.7$ nM for CXCR1 was reported.[55] Several other CXCL8 mutants with mutations at the dimer interface have been described: L25Y/V27R, being still monomeric at 10 μM, E24L/L25E, V27A,[56,57] as well as L25F/V27F with a dimerisation constant of $K_d = 7$ μM.[58] Furthermore, a shortened variant of CXCL8, CXCL8(1-66), was reported to be monomeric and characterized by ultracentrifugation, nuclear magnetic resonance (NMR) and circular dichroism (CD) spectroscopy.[59,60]

Another important aspect of CXCL8 function is the binding to gylcosaminoglycans (GAG) on the endothelium in order to form stable haptotactic gradients *in vivo*.[22,61–63] Heparin has been used as the representative GAG in studies in the literature, because it is more easily obtained and cheaper than other GAG and it has a similar structure to heparan sulfate, the most abundant GAG that is expressed on 'virtually every cell in the body'.[64,65] Interaction with GAG is a feature of many different proteins and not restricted to the chemokine family. Consensus sequences for GAG-interaction, e. g. XBBXBX or XBBBXXBX, where B is 'the probability of a basic residue and X is a hydropathic residue', have been identified in several proteins.[66] Protein:GAG interactions largely depend on the type and spacing of basic residues and the affinity depends largely on the arginine to lysine ratio in positions that pair with GAG anions.[67] Hydrophobic interactions between N-acetylated groups have also been suggested to play a role in specific binding to heparan sulfate.[67] The heparin:CXCL8 interaction is mediated by basic residues K67, K64, R68 and R60 on the C-terminal α-helix of CXCL8.[22–24] When the α-helix is missing, CXCL8 can no longer bind to GAGs on endothelial cells and there is no consecutive internalization and transcytosis of CXCL8 across endothelial cells as shown in experiments with rabbit skin.[24]

3.1.2 Leukocyte Extravasation

Pattern-recognition receptors (PRRs) can bind to specific pathogen-associated molecular patterns (PAMPs) to ensure the detection of invading microorganisms.[68] The recognition of pathogens in the human body is ensured by a class of PRR called toll-like-receptors (TLR) that are expressed by endothelial cells.[69] Bacterial proteoglycans or lipopolysaccharides (LPS) bind to a plasma-protein such as lipopolysaccharide-binding protein (LBP). This complex is recognized by the phagocytes via binding to the cell-surface protein CD14 (cluster of differentiation 14). This protein complex is capable of activating TLR4 which induces a signalling pathway that eventually causes inflammation by activating DNA-promotors for the transcription of genes that code for pro-inflammatory proteins like chemokines that guide white blood cells to the site of inflammation (see Figure 3.2).[70,71] To reach the inflamed tissue, leukocytes have to leave the blood vessel across the endothelial border ('extravasation' or 'diapedesis'). An important factor in the process of leukocyte extravasation is the interaction between leukocytes and endothelial cells in the blood vessel. Apart from chemokines also selectins and integrins play a key role in the 'leukocyte adhesion cascade' that is characterised by the following steps: leukocyte tethering, rolling, activation, arrest, intravascular crawling and transmigration.[72]

Leukocyte tethering and rolling is mainly dependent on selectins. Selectins are a class of sugar-binding transmembrane glycoproteins with a highly conserved extracellular structure.[73] The extracellular N-terminus consists of a sugar-binding lectin-domain, an EGF-like (epidermal growth factor-like) domain and a number of consensus repeats of complement regulatory protein (CR) homologous domains depending on the type of selectin: two repeats for L-selectins (*l*eukocyte), six repeats for E-selectins (*e*ndothelium) and nine repeats for P-selectins (*p*latelet).[74] Inflamed vascular endothelial cells express increased amounts of P- and E-selectins which are able to bind P- and E-selectin glycoprotein ligand-1 (PSGL-1, ESGL-1) on the membrane surface of leukocytes with a calculated P-selectin:PSGL-1 dissociation constant between $K_d = 3 \pm 2$ nM and $K_d = 320 \pm 20$ nM.[75,76] The binding strength is dependent on shear stress which means that the blood flow strengthens the bond ('catch-bond') thereby facilitating the capture, i. e. teth-

ering, of the leukocyte. The binding between the lectin-domain and the glycosylated protein is reversible which enables the leukocyte to perform a rolling-motion on the vascular endothelium.[72,77–79]

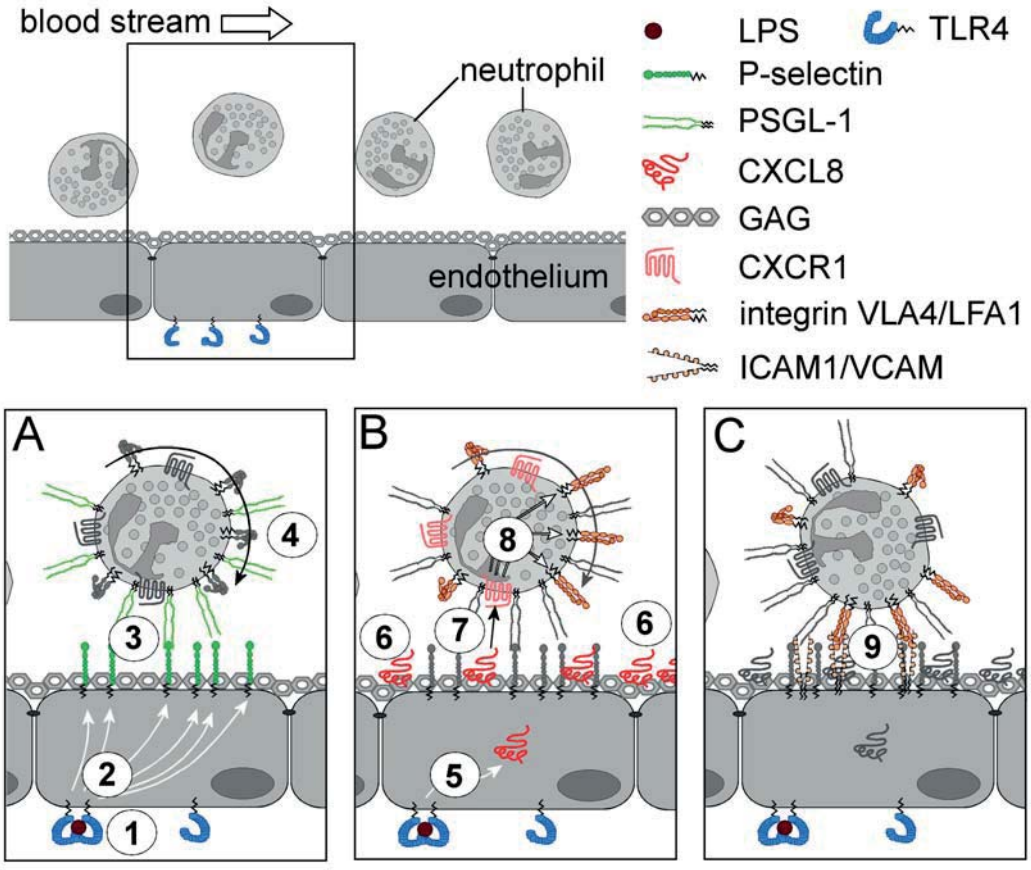

Figure 3.2: The involvement of chemokines in the extravasation of leukocytes. Neutrophils regularly circulate the blood. (A) When a foreign substance enters the body (e. g. a bacterial lipopolysaccharide LPS), it is detected by binding to a receptor ① (e. g. a toll-like receptor TLR4). This initiates a signalling cascade ② that leads to an increased expression of P-selectins. ③ P-selectins bind to P-selectin binding glycoprotein ligand PSGL-1 on neutrophils. This interaction is dependent on shear stress ('catch bond') and leads to a rolling motion ④ of the neutrophil on the endothelium. Leukocyte rolling is also observed in the absence of pathogens, since P-selectins are expressed regularly on the endothelium. The effect is however increased by enhanced P-selectin expression. (B) The TLR-4 activation also leads to the expression of chemokines ⑤ (e. g. CXCL8) that are presented on the endothelial glycosaminoglycans (GAG) ⑥ in the blood vessel. ⑦ Chemokine CXCL8 activates the G-protein coupled receptor (GPCR) CXCR1. This interactions leads to the activation of integrins (e. g. very late antigen 4 VLA4 and lymphocyte function-associated antigen 1 LFA1) ⑧ on the neutrophil (switchblade-like movement, affinity regulation) most likely via the expression of talin-1 and/or kindlin-3. (C) The firm adhesion of the neutrophil to the endothelium due to the interaction of the integrins with cell-surface glycoproteins like intercellular adhesion molecule 1 (ICAM1) or vascular cell adhesion molecule (VCAM1) ⑨ leads to the arrest of the neutrophil on the endothelium. The neutrophil then extravasates into the tissue where it is guided to the site of inflammation by a chemokine gradient that may be stabilized by CXCL8:GAG interactions.

The arrest and firm attachment of leukocytes to the endothelium is mediated by integrins. Integrins are transmembrane receptor proteins that enable cells to attach to extracellular matrices (e. g. fibrinogen,

fibronectin) or cells (via integrin-binding cell adhesion molecules, CAMs).[73] They are non-covalently associated heterodimers consisting of α- and β-subunits which both possess one transmembrane region.[80] They combine to form the integrin structure via association of their amino-terminal extracellular loop domains. When chemokines are secreted during inflammation, they activate integrins on leukocytes through G-Protein coupled receptor-dependent signalling: Chemokines bind to their GPCR thereby activating several different pathways such as the PI3K, the RhoA and RAP1 pathway ultimately leading to the enhanced production of talin, a 250 kDa cytosceletal protein with a C-terminal actin-binding domain and an N-terminal region for integrin-binding. Talin binds the intracellular parts of the integrin β-chain, leading to a conformational change.[81] This switchblade-like movement of the integrin head, which initially is bent towards the membrane, increases the affinity for integrin ligands (affinity regulation) such as intercellular adhesion molecule (ICAM1) or vascular cell adhesion molecule (VCAM1). The binding strength may be further enhanced by the interaction of several integrins on a localized spot of the plasma membrane, the so called integrin-clustering (valency regulation).[81–83] The activation of integrins through GPCR signalling is referred to as 'inside-out signalling'.[84,85] The neutrophil, being firmly attached to the endothelium, may then start to crawl on the endothelium. This crawling on the endothelial surface enabled by the integrin:CAM interaction is referred to as two-dimensional 'haptokinesis'.[3] Neutrophil:VCAM interactions cause an increase in intracellular calcium ions that leads to the phosphorylation of cadherins (calcium dependent glycoproteins) which causes loosening of cell-cell interactions.[6] Neutrophils may then extravasate between endothelial cells and, guided by a chemokine gradient, migrate into underlying tissue disrupting the extracellular matrix by release of enzymes like elastase.[86]

3.1.3 Cell Migration

The two dimensional migration of white blood cells on the endothelium is directed towards an increasing amount of chemoattractant and may therefore be classified as 'chemotaxis' (if the chemoattractant is dissolved in the surrounding fluid) or 'haptotaxis' (if the chemoattractant is immobilized on a surface, e. g. chemokines bound to GAG on the cell surface). Cell migration is dependent on the rearrangement of the cytoskeleton, namely the rapid polymerisation and depolymerization of actin. Actin is one of the most abundant proteins in eukaryotic cells.[87] Filamentous F-actin is formed by the association of globular G-actin enabled by rotation of the two G-actin subunits relative to each other.[88,89] All directed cellular movements rely on motor proteins like myosin that bind to and move along cytoskeletal filaments like actin.[89,90] Myosin (non-muscle myosin II) movement along actin filaments is based on a conformational change of the actin-binding head referred to as 'power stroke' that requires the hydrolysis of adenosine triphosphate (ATP) to adenosine diphosphate (ADP). Myosin therefore has ATPase activity. In association with myosin, electron microscopy images of actin filaments look like the close follow-up of arrowheads, so one end is called the pointed or (-) end and the other is called barbed or (+) end.[87] Actin polymerization is diffusion-dependent, i. e. dependent on the rate of G-actin/F-actin collision at the (+) end of the filament. Rate constants for ADP- and ATP-actin association and dissociation vary at the (+) and (-) ends of the filament. The – compared to ADP-G-actin – higher association rates of ATP-G-actin at the (+) end together with the high dissociation rates of ADP-G-actin at the (-) end lead to a 'treadmilling' of G-actin monomers from the (+) to the (-) end.[87] The rate of polymerisation is thus dependent on the amount of monomeric ATP-G-actin available in the cell as well as the availability of barbed (+) ends for polymeriza-

tion. Regulatory proteins like profilin and thymosin-β4 help to keep the amount of G-actin high: profilin binds one side of G-actin and thus suppresses the association of G-actin at the (+) end. Thymosin-β4 stops actin polymerization altogether and competes with profilin for the binding of G-actin.[87] Capping proteins prevent further polymerization at the barbed (+) ends. F-actin polymerization in cells can be studied by phalloidin-staining. Phalloidin is a toxic bicyclic heptapeptide isolated from the death cap fungus (*Amanita phalloides*) that binds F-actin and prevents filament depolymerization.[91,92] When treated with phalloidin labelled with a fluorophore, the amount of filamentous actin can be directly measured by means of fluorescence intensity.[92] Cells are fixed with formaldehyde at different timepoints seconds after stimulation with a chemoattractant and consecutively stained to visualize the effect of increased actin-polymerization by binding of fluorescently labelled phalloidin.[93,94]

Amoeboid movement of neutrophils starts with the polarization of the cell, i. e. the change of cell morphology that leads to an asymmetric, flattened cell shape that makes it possible to distinguish cell front and rear (see Figure 3.3).[95–97] The frontal structure is characterized by the lamellum, a flat protruding cell structure of finger-like, outward pointing actin filaments.[98] Nucleus and other organelles are concentrated at the rear, the so-called uropod. The outer rim of the lamellum, the lamellipodium, is characterized by an increased amount of actin-polymerization with Arp2/3 crosslinked, myosin-free actin filaments ('dendritic actin'). Behind the outer lamellipodium, the lamellum is characterized by thick bundles of actin, myosin and actin-binding proteins.[99] Polarization is followed by protrusion: actin filaments push the membrane forward by polymerizing towards it (elastic Brownian ratchet model) thus forming pseudopods.[100,101] The force of the F-actin pushing towards the membrane also leads to a backwards push of the filaments towards the uropod, also termed 'retrograde actin flow', characteristic for the lamellum.[102] Focal adhesions (e. g. by the interaction of actin filaments with integrins, mediated by proteins like talin) between cell and substrate can provide points of force transfer enabling a faster migration (integrin clutch hypothesis).[102,103] Myosin contracts the filamentous network in the lamellum and with the retraction of the uropod the cell moves forward.[102]

The role of CXCL8 signalling in actin polymerization, pseudopod formation and ultimately amoeboid movement is complex and not understood in detail. Early studies showed the CXCL8-induced calcium release in neutrophils is sensitive to pertussis toxin.[110] CXCL8 receptor CXCR1 was thus believed to associate with $G_{i\alpha}$ (*inhibitory*) G-proteins.[111,112] The increase in calcium can be explained by the G-protein βγ subunit activating phospholipase Cβ of the IP_3 (inositol-1,4,5-triphosphate) signalling pathway.[113–115] Calcium ions are however not responsible for actin polymerization but rather facilitate actin depolymerization.[116] More recent studies found that CXCR1 may also associate with G-protein $G_{\alpha 16}$.[117] CXCL8 increases phosphatidyl-inositol-3,4,5-triphosphate (PIP_3) content in human neutrophils through PI3K activation; inhibition of PIP_3 formation results in reduced rate of pseudopod formation upon CXCL8 stimulation.[107] PIP_3 activates protein kinase B (Akt/PKB) and leads to the activation of several proteins involved in actin polymerization and cytoskeleton rearragement (e. g. myosin light chain kinase MYLK for cell spreading). Small GTPases like Rho, Rac and Cdc42 are major factors in rearrangement of the cytoskeleton.[118] Cdc42 for instance leads the activation of Arp2/3, a protein that is responsible for actin crosslinking to enhance filament stability.[107–109]

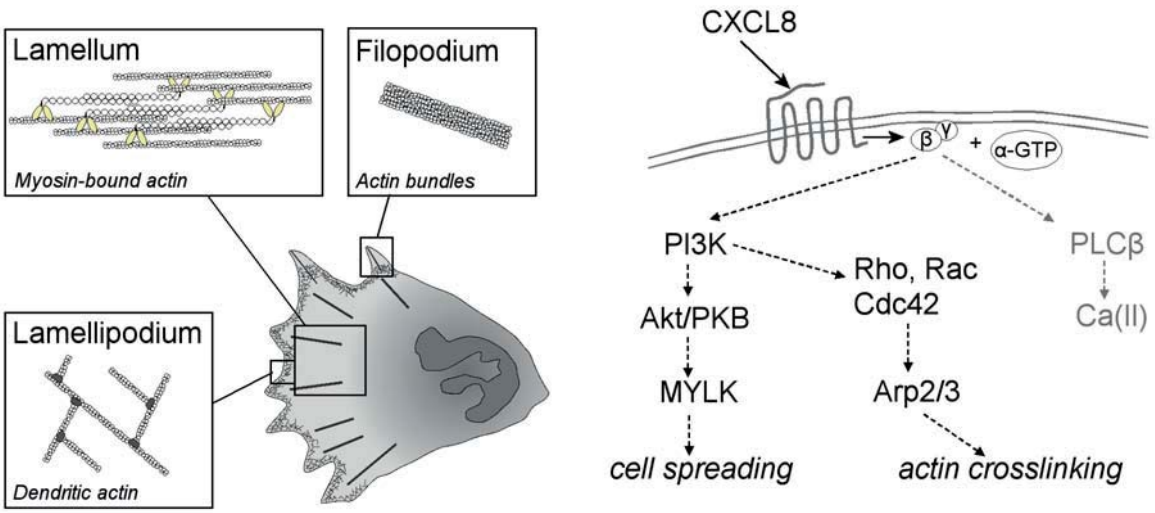

Figure 3.3: Actin polymerization in human neutrophils in the process of cell migration. The amoeboid movement during migration depends on restructuring of the cytoskeleton, especially the actin filaments in the cell. Stimulated cells polarize by developing a flat outstretching structure in the front. This so-called lamellum is characterized by myosin-bound actin bundles that enable cell contraction. At the outermost rim of the cell the lamellipodium, a tight network of short, Arp2/3 crosslinked myosin-free actin filaments ('dendritic actin') allows for rapid actin reorganization and changes in movement direction. Filopodia of stable actin bundles are pushed outwards to scan the environment and enable protrusion of the cell.[99, 102, 104] CXCL8 plays an important role in actin polymerization but only parts of the signalling cascade are known.[105] After GPCR-activation, the G-protein βγ subunit activates phosphatidyl-inositol-3-kinases (PI3K) that phosphorylate PIP2 to PIP3, an important factor in neutrophil chemotaxis.[106, 107] Further down the cascade, protein kinase B (Akt/PKB) is activated and may activate myosin light chain kinase (MYLK), a process likely to be important for cell spreading.[108] Small GTPases like Cdc42 are also activated by PI3K leading to the activation of Arp2/3, the protein that enables the crosslinking of actin filaments in the lamellipodium.[107–109] The pathway of phospholipase Cβ leading to the formation of inositol-3,4,5-triphosphate (IP3) and ultimately to the release of calcium ions into the cytosol is also activated, the role of calcium for cell migration is not entirely clear though.

3.1.4 Chemokines and Diseases

Inflammation-stimulating (proinflammatory) chemokines like CXCL8 provoke an immune responses of the host system. This initial response is primarily focussed on the destruction of invading pathogens. When recruited, neutrophils can release aggressive substances like reactive oxygen species (ROS) or proteases and other enzymes to ensure pathogen destruction. Wrongly induced or persistent inflammation however may lead to autoimmunity and chronic inflammation and thus to unwanted damage of the host organism. Examples for chemokine-related diseases are autoimmune disorders like multiple sclerosis and rheumatoid arthritis and inflammatory diseases like asthma.

Multiple sclerosis (MS) is a degenerative disease of the nervous system where the nerve-protecting myelin sheath is attacked by the immune system. Elevated levels of CCL3 (MIP1α), CCL4 (MIP-1β) and CCL5 (RANTES) were found in the affected tissue of MS patients. Also high levels of CXCL10 (IP-10) and abnormally low levels of CCL2 (MCP-1) were detected in the cerebrospinal fluid,[119–121] though an even

greater number of chemokines may be involved in the disease.[122] The etiology of MS or even details about how chemokines influence the development of the disease are unknown. One hypothesis suggests that chemokines induce immune cells to cross the blood-brain barrier, a prerequisite for attacking cells in the central nervous system.[123]

In rheumathroid arthritis an unknown cause leads to the prolonged inflammation of the joints with progressive cartilage and bone damage. Chemokines are thought to be involved in the extravasation of leukocytes through the vascular endothelium into the synovium and the persistent recruitment of macrophages into the joint.[124] It is hypothesized that an imbalance in the complexly regulated system of cytokines with a multitude of different receptors and ligands/effectors has an influence on cell differentiation.[124] Especially chemokines CCL5 (RANTES), CCL3 (MIP-1α) and CCL2 (MCP-1) are believed to be a critical factor in the proliferation of T-cells.[125,126]

Apart from the production pool in the bone marrow, there are three marginated pools of neutrophils in the body: lung, liver and spleen each contain an elevated amount of neutrophils due to the prolonged passing time (margination time) through these organs.[2] Neutrophils pass most slowly through the lung making it prone to diseases related to damage caused by neutrophil attacks. Asthma is broadly defined as a reduced function of the lung due to allergens causing prolonged inflammation that leads to severe airway damage. Eosinophils have been identified as the cell type mainly involved in the disease. Chemokines like eotaxin induce the migration of eosinophils and are therefore implicated in the process of lung inflammation. In a mouse-model it has been proven that eotaxin-1 (CCL11) and eotaxin-2 (CCL24) both play a crucial role in recruiting eosinophils into the lung.[127] The high levels of eosinophilia were lost in mice lacking the eotaxin receptor CCR3. The eotaxin/CCR3 pathway is therefore likely to play a major role in the development of the disease.

An increased amount of chemokine receptors in cancer cells suggests an involvement of chemokines in cancer development.[128] The angiogenic nature of chemokines that possess the ELR motif (glutamic acid - leucine - arginine sequence close to the N-terminus) has been identified from early on.[13] Chemokines can thus play a crucial role in cancer development not only regarding the migration of cancer cells that leads to metastasis but also the development of new blood vessels that supply a tumor with oxygen and nutrients. A correlation between the expression of CCR7 and the aggressiveness of lung-cancer has been described.[129] Another well-discussed tumor survival strategy is autocrine signalling, the production of growth factors, chemokines and other proteins by the tumor cells themselves.[130] Many chemokines were found to play a role in tumor autocrine signalling *in vitro*: CXCL8 for example has been implicated in the growth of human skin cancer melanoma and colon, liver as well as pancreatic cancer cells.[131–133]

Another prominent example for the involvement of chemokines in diseases is the infection of human immune cells by the human immunodeficiency virus (HIV) that leads to the acquired immunodeficiency syndrome (AIDS). A viral glycoprotein-complex binds the CD4 receptor on human T-helper cells or macrophages followed by a conformational change in the glycoprotein-complex and the subsequent binding of a co-receptor, usually of the chemokine class, like CXCR4 or CCR5.[134] Eventually the viral membrane fuses with the membrane of the cell thereby releasing the viral DNA into the cell.[135] Monocyte chemoattractant protein-1 CCL2 (MCP-1) levels in the plasma were found to correlate with the virus titer in HIV.[136] Furthermore, in a tissue culture of the human blood-brain-barrier, the stimulation with CCL2 lead to the transmigration of

HIV-infected leukocytes across the barrier.[137] This way, the virus could enter the central nervous system and cause NeuroAIDS or HIV-associtated dementia by attacking and destroying brain tissue.

3.1.5 Chemokine Inhibitors

Since chemokines are affiliated with a number of different severe diseases (see previous Section) there has been a keen interest in finding inhibitors of chemokine functions during the past two decades. There are two obvious approaches for inhibiting the effect of chemokines: the inhibition of the receptor of the respective chemokine or the inhibition of the chemokine itself. The chemokine network is very complex with several chemokines binding to the same receptors. The inhibition of a receptor therefore has an effect not only on the targeted chemokine, but also on all other chemokines interacting with the receptor. In order to suppress the function of a single chemokine, it is therefore favourable to target the chemokine itself. This section therefore focuses on inhibitors for chemokine ligands. Several different classes of chemokine inhibitors are known: peptides, small molecules, proteins. Also several antibodies against various chemokines are commercially available.

As early as 1993 the first peptides capable of inhibiting CXCL8 were studied. It was found, that the N-terminus of CXCL8, CXCL8(3-25) was capable of inhibiting the binding of CXCL8 to human neutrophils and inhibited the chemotactic activity of recombinant human CXCL8.[138] The short peptide Ac-RRWWCR-NH_2, later named 'antileukinate', was discovered through the screening of a synthetic peptide combinatorial library of mixtures of hexapeptides.[139] The sequence was initially tested for antimicrobial activity against *Staphylococcus aureus* but was later found to inhibit the binding of CXCL8 to human neutrophils and thus also inhibit CXCL8-induced neutrophil chemotaxis.[140] Further studies revealed that the peptide also suppresses the staphylococcal enterotoxin A (SEA)-induced recruitment of neutrophils into the lung of rabbits and that it is capable of preventing CXCL1 from binding to human melanoma cells.[141–143]

Peptides like Ac-MWDFDD-*spacer*-MPPADEDYSP-NH2 which are derived from the extracellular N-terminus of CXCR1, one of the binding sites for CXCL8, bind to CXCL8 and possess K_i values in the lower micromolar range.[144,145] The complex of one of these peptides, CXCR1-p1, i. e. CXCR1(9-14) linked to CXCR1(20-29) by 6-aminohexanoic acid (sequence: MWDFDD-Ahx-MPPADEDYSP, K_i =7 μM), with CXCL8 was further studied by NMR (pdb id: 1ILQ).[42] Experiments revealed that CXCR1-p1 is unstructured (see Figure 3.4) and retains the previously confirmed[26,27] interaction sites of CXCL8 with the N-terminus of CXCR1: N-loop, 3_{10} helix, 40S loop, β_3 strand. In a similar approach, peptides derived from the N-terminal domain of CCR3 have been found to be potent CCL11 (eotaxin) inhibitors while peptides based on the sequence of the first and third extracellular loops showed no binding capacity.[146] Peptides derived from the N-terminus of rabbit CXCR1 have also been reported.[147] In 2009 consensus sequences from several different chemokine receptors were presented on recombinant phage and screened on microarrays for their binding affinities to 31 different chemokines.[148] The rationale was to find low-affinity chemokine-binding peptides that are capable of simultaneously modulating the function of several chemokines in the chemokine network. Peptides LFGNDCE and WVFGNAMCK were found to be active in delayed type contact hypersensitivity (DTH) in mice. Phage display has also been employed for the discovery of high affinity chemokine inhibitors: the peptide CPWYFWPC has been shown to inhibit the interaction of CCL11 with its receptor CCR3, and SAMWDF as well as FWDDFW were found to interfere with CXCL8:receptor interaction thus inhibiting CXCL8-induced neutrophil migration.[149,150] It is however

not specified, if these peptides bind to the chemokine or the receptor. CROSS proteins (chemokine receptor elements on a soluble scaffold) were designed to study the interactions of chemokines with their receptors, by fusing chemokine receptor sequence parts onto a protein scaffold (B1 domain of streptococcal protein G) that retains the spatial orientation of parts in the native receptor.[44,151] They have not been tested for inhibitory effects but they have been found to bind to chemokines with affinities in the low micromolar range.

The first small molecules known to bind CXCL8 were (*R*)- and (*S*)-ketoprofene (see Figure 3.4). They suppress the chemotactic activity of human monocytes and human polymorphonuclear neutrophils *in vitro*.[152,153] Chalcones (see Figure 3.4), similar structures derived from 1,3-diphenylurea reduce the binding of CXCL12 to CXCR4, inhibit calcium response, prevent receptor internalization in human kidney cells and block the recruitment of eosinophils into the airways of ovalbumin-challenged mice *in vivo*.[154,155]

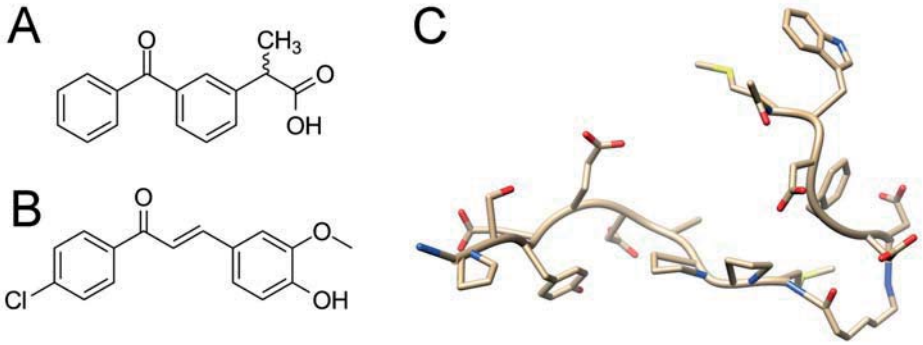

Figure 3.4: Examples of known chemokine inhibitors. (A) Ketoprofene isomers bind to CXCL8 and inhibit neutrophil migration.[152,153] (B) Chalcone 4 binds CXCL12 and inhibits the interaction with its receptor CXCR4.[154] (C) NMR-based structure of CXCR1-p1 (pdb id: 1ILQ), a peptide derived from the N-terminal sequence of CXCR1 that binds CXCL8.[42,145]

NOXXON pharma introduced D-RNA aptamers with promising binding capabilities that were re-synthesized as L-RNA oligonucleotides and named 'spiegelmers' ('mirrored', German: *'spiegeln'*). They are stable against nucleases and therefore could serve as potential drugs. An inhibitor of CCL2 (MCP-1) is commercially available and proved to be effective in the treatment of lupus nephritis in mice.[156] A commercially available anti-CXCL12 (SDF1-α) spiegelmer proved to effectively reduce the migration of chronic lymphocytic leukemia cells in Transwell assays.[157] A broad-range aminolactam chemokine inhibitor (FX125L) is being tested in clinical trials mainly for the treatment of asthma.[158]

To evade immune responses, some organisms in nature prevent the release of or inhibit the function of chemokines in order to be able to live on the host. Tick saliva is known to contain small proteins called 'evasins', which show chemokine neutralizing activity: Evasin-1 for instance is specific for CCL3, CCL4 and CCL12 and binds to these chemokines with an affinity in the low to sub-nanomolar range.[159] Large DNA viruses like the herpes- and poxviruses produce chemokine-mimicking proteins as well as chemokine-receptor mimics and chemokine-binding proteins that prevent the interaction of chemokines with their receptors.[160,161] Myxoma virus T7-glycoprotein binds the GAG-binding domain of CXCL8 and is believed to thus interfere with neutrophil recruitment.[162]

Proteins are involved in all types of cellular processes, e. g. signalling and detection, catalysis and reproduction. All these reactions depend on the interaction of different classes of proteins like receptors and signalling proteins, enzymes and cofactors. Influencing the interactions of proteins can have desirable effects for the host organism and lead to the development of new drugs. Protein-binding substances, or 'ligands', can be found in many substance classes: small molecules, peptides/proteins, peptidomimetics or nucleic acids.

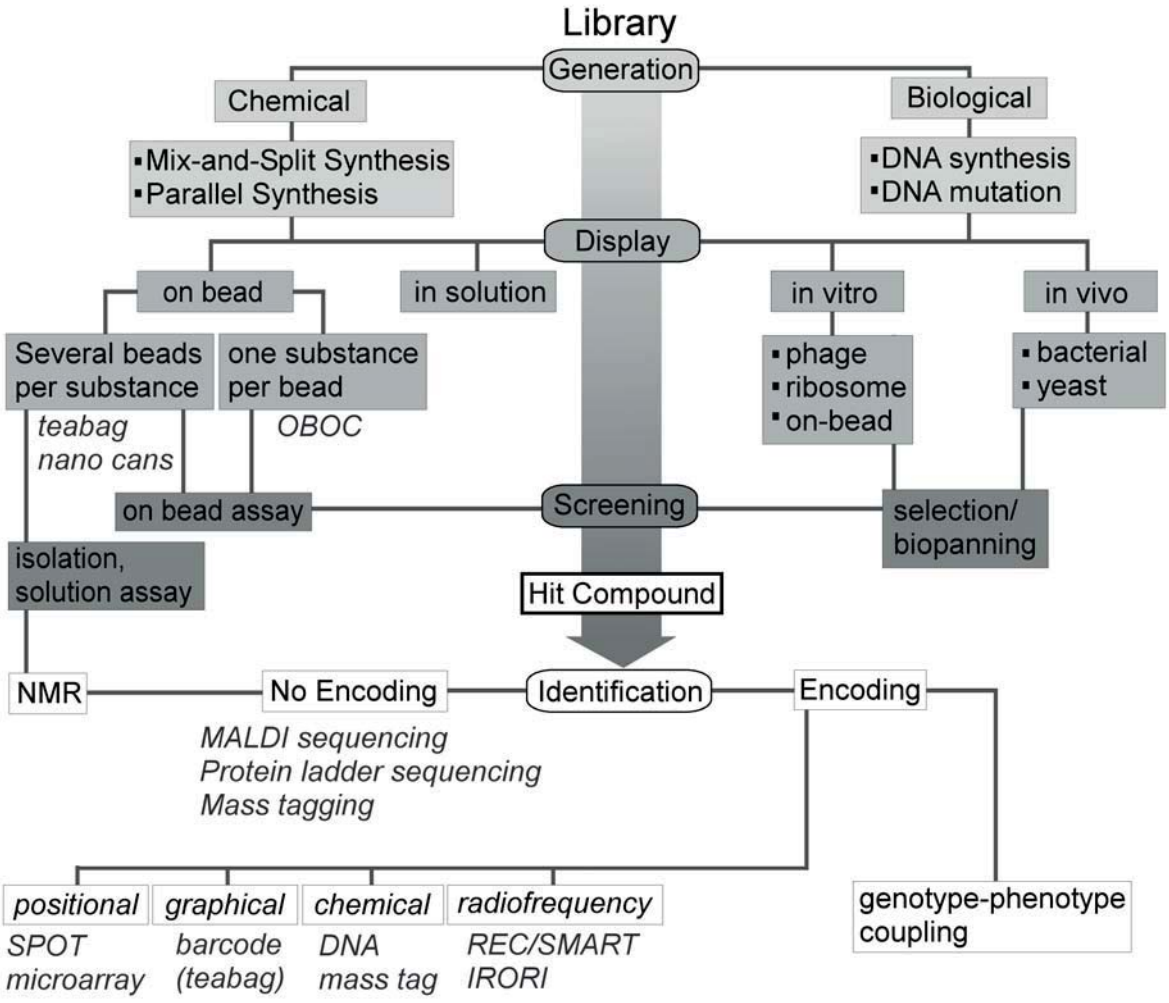

Figure 3.5: Finding novel protein ligands by library screening: Overview of different methods for library generation, display, screening and hit identification. Libraries are generated by chemical synthesis (peptides, DNA) or mutation (DNA, RNA) to give 'chemical' and 'biological' libraries. They are displayed on bead, on phage or on microorganisms. Different types of libraries are screened by testing compounds in *in vitro* or *in vivo* assays that produce a readout (like fluorescence intensity) for each compound, permitting to find active compounds (hits). Hits are then identified by sequence deconvolution or their assay-specific encoding. Chemical, positional, graphical and radiofrequency encoding have been described for chemical libraries.[163] Biological libraries are characterized by their genotype-phenotype encoding. Identified hits may be re-tested and further characterized for their biochemical properties.

The discovery of novel protein ligands either relies on rational ligand design or on the synthesis of large compound collections (libraries) that are tested (screened) in biochemical assays to find compounds with certain desired properties. Methods for the synthesis as well as for the screening of compound libraries are diverse. An experimental approach to studying protein-protein interactions is the parallel or combinatorial synthesis of a large number of potential binding molecules and a subsequent assay for the visualization of a positively interacting substance. The use of a large number of substrates increases the possibility of finding a specific protein ligand and minimizes the problem of synthetic bias that occurs when synthesizing and testing only a small number of previously chosen substances. To this day several different types of libraries with different kinds of protein-targeting molecules are known. They are generally divided into synthetic (or chemical) and biological libraries. Chemical library synthesis relies on the concept of combinatorial chemistry: the generation of compounds by repetitive or combined use of well-characterized, high-yield chemical reactions (like amide-coupling or click-chemistry, respectively). Biological libraries on the other hand are created by the use of microorganisms for peptide synthesis, e. g. the synthesis and presentation of peptides on phages in a phage display library.
The screening of compounds in chemical and biological libraries usually comprises an *in vitro* or *in vivo* assay that allows a simple readout (e. g. by fluorescence intensity or luminescence) to distinguish between binding or active compounds and non-binding or inactive compounds. Promising candidates identified by the readout are termed 'hits' or 'screening positives'. The chemical identity or sequence of the hits is then identified depending on the mode of library encoding. Hits may then be re-synthezised and further characterized for affinity and activity. The different types of libraries and encoding methods are described more closely in the following paragraphs and are summarized in Figure 3.5.

3.2.1 Chemical Libraries

Most chemical library synthesis reactions are carried out on solid support. Solid phase synthesis was introduced by Merrifield in 1963 for the synthesis of peptides on chloromethylated crosslinked polystyrene.[164] In general the attachment to solid support ensures an easy removal of unattached reaction side products after each synthesis step, enables the use of chemicals in excess and facilitates the automation of the synthesis which can provide a large number of substances in small or large quantities. Limitations of solid-phase synthesis are the inability to perform purification steps to remove side-products like truncated or erratic sequences during the synthesis and the consecutive analysis of small quantities of compounds. For the synthesis and screening of large libraries the question of compound analysis is essential. There are different approaches to overcome the problem of direct substance identification: Several kinds of library encoding methods have been reported, like positional, chemical, graphical and radiofrequency encoding.[163]

Non-Encoded Libraries

The basic variant of combinatorial libraries is non-encoded: various compounds are synthesized in a combinatorial process, and important sequences are only identified after positive screening results. This determination of compound identity is also termed 'deconvolution'. At first, the synthesis of compounds and the presentation as an array for screening were two seperate steps: compounds were synthesized and then presented on a solid support for screening ('spotted' onto a surface). Sometimes the term 'probe' is used for

a spotted compound with a known identity, while 'target' describes an unknown compound in an array.[165] The dot blot is an early example for pre-synthesized DNA probes spotted onto nitrocellulose filters in an array format.[166] When the printed spots became smaller and the compound collections were printed on a minimal space, the term 'microarray' was introduced: a densely printed collection of biomolecules on surfaces (biochips) for screening purposes,[167] e. g. microarrays of DNA,[168] peptides[169] and natural products.[170] These techniques link a compound to a certain space on an array, but do not directly identify the sequence, which has to be determined by further analysis.

The synthesis and presentation of the library can also be combined in one process. A combinatorial process known as the 'mix-and-split' (also 'mix-and-pool') method was reported by Furka in 1991.[171] It is especially useful for the synthesis of polymeric, repetitive structures like peptides: the first amino acid is coupled to the resin and the resin is split into several portions, each portion is reacted with a different amino acid and the resin is pooled and mixed. After splitting, the next set of amino acids is attached and so on (see Figure 3.6). The number of possible peptides N is dependent on the amount of amino acids A_i per cycle i and the number of residues of the peptide n (equals the total number of synthesis cycles):

$$N_n = \prod_{i=1}^{n} A_i \tag{3.1}$$

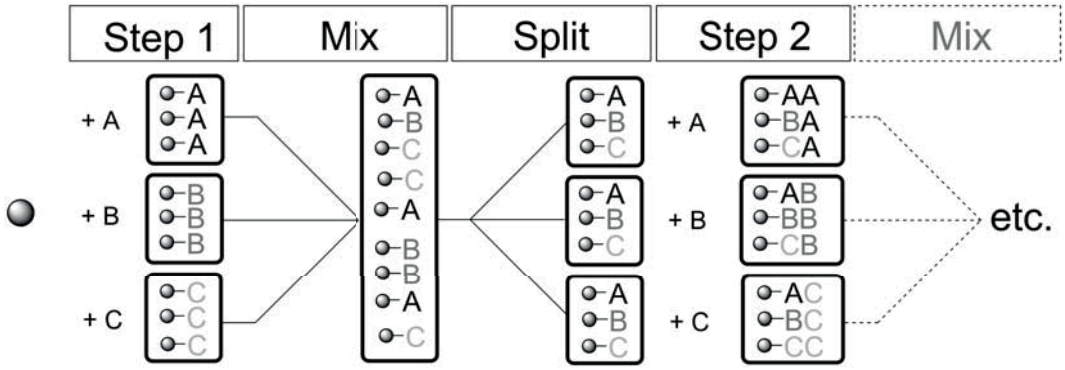

Figure 3.6: Schematic explanation of mix-and-split synthesis. One synthesis cycle consists of reaction step, mixing and splitting of the reactants. The number of compounds synthesized increases exponentially with the number of synthesis cycles.

The mix-and-split method was further developed by Lam to produce libraries in which each individual resin bead carries one sequence, the so-called 'one-bead-one-compound' (OBOC) libraries.[172] Compounds from this type of library can be easily screened 'on-bead': proteins are incubated with the library and give a certain signal readout, e. g. by fluorescence intensity of an attached fluorophore. OBOC libraries with various non-peptide oligomers and small molecules have been reported and have been reviewed, e. g. by Lam *et al.*[173] The sequence deconvolution relies on classical fragmentation or sequencing methods like matrix-assisted laser desorption/ionization time-of-flight mass spectrometry with ion fragmentation (MALDI TOF MS/MS) or Edman degradation. The analysis is therefore largely restricted to peptides, peptidomimetics and nucleic acids. The OBOC method needs no additional synthesis steps and no further specialized decoding instrumentation and thus keeps synthesis and analysis prizes low. Mass spectrometric analysis of unlabeled peptide or peptoid OBOC libraries may be facilitated by protein ladder sequencing.[174,175]