TECHNISCHE UNIVERSITÄT MÜNCHEN

Lehrstuhl für Biochemie

Structural And Biophysical Characterization Of Antibody:Antigen Complexes Of Therapeutic Relevance

Felix Ludwig Schiele

Vollständiger Abdruck der von der Fakultät für Chemie der Technischen Universität München zur Erlangung des akademsichen Grades einer Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigten Dissertation.

Vorsitzender:		Univ.-Prof. Dr. Johannes Buchner
Prüfer der Dissertation:	1.	Univ.-Prof. Dr. Michael Groll
	2.	Univ.-Prof. Dr. Thomas Kiefhaber

Die Dissertation wurde am 07. 02. 2013 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 12. 04. 2013 angenommen.

Contents

List of Tables

List of Figures

1 Introduction

1.1 Antibodies

1.1.1 Immune System and Generation of Antibodies

The Innate Immune System The immune system in vertebrates provides a defense mechanism against foreign molecules such as viruses and bacteria. When the body's physical barriers like skin or mucosae are overcome, the innate immune system reacts on broad classes of pathogens that have certain molecular patterns immediately after infection (Aderem and Underhill, 1999). It depends solely on preformed receptors which are already encoded in innate germline cells. One example are dendritic cells (resident tissue macrophages) that are derived from circulating monocytes which express CD14, a receptor for bacterial lipopolysaccharide. Upon binding of a pathogen, the receptors cluster on the surface of the dendritic cell which induces phagocytosis and subsequent lysosomal degradation (Gordon, 1998). The resident macrophages are also capable of releasing cytokines and chemokines, which recruit immune cells like the phagocytic neutrophils to the site of infection. Other cells that are activated by the innate immune system are natural killer cells, mast cells, eosinophils, basophils and macrophages (Brightbill et al., 1999).

Besides this cell-mediated response, the innate immune system is able to induce activation of the complement system. It comprises several proteins including proteases that are themselves activated by proteolytic cleavage and protects against infection in three ways (Janeway et al., 2001): i) In a process called opsonization, large numbers of activated complement proteins bind covalently to pathogens. Phagocytes with receptors specific for the complement proteins can subsequently engulf and eliminate the marked pathogens (Gasque et al., 2004). ii) Fragments of complement proteins can induce inflammatory response, resulting in vasodilatation of blood vessels and increasing vascular permeability. These fragments also act as chemokines that recruit additional inflammatory phagocytic cells, like neutrophiles and monocytes to the pathogen (Schraufstatter et al., 2002). iii) Finally, proteins of the complement system are capable of forming a membrane-attack complex. This multi protein complex inserts into bacterial membranes and adopts a channel-like structure. It disrupts the lipid bilayer and allows for free diffusion of water and solute molecules which results in the loss of cell homeostasis and subsequent destruction of bacteria (Peitsch and Tschopp, 1991).

The Adaptive Immune System When the innate immune response is bypassed or evaded by a pathogen, the adaptive immune response is activated. In contrast to the

innate immune system, adaptive responses are highly specific to the pathogen that induced them. A substance that is able to induce adaptive immune response and thereby the production of antibodies (also referred to as immunoglobulins (Tiselius and Kabat, 1939)) is called antigen (antibody generator). Antigens can bind to specialized lymphocytes, B-cells, that express a membrane-anchored antibody molecule (B-cell receptor). The binding of the antigen to the B-cell receptor, with the aid of a helper T-cell, stimulates the proliferation and differentiation of the B-cell into an antibody-secreting effector cell (plasma cell). These cells produce large amounts of specific antibodies that have the same antigen-binding specificity as the B-cell receptor of the originating B-cell (Alberts et al., 2002).

1.1.2 Structure of Antibodies

All antibodies share a common structural unit that is composed of four polypeptide chains, two identical light chains (LC) of about 220 amino acids and two identical heavy chains (HC) of about 440 amino acids (the number of amino acids depends on the antibody class, see below). The four chains interact via noncovalent forces and covalent disulfide bonds (Edelman and Poulik, 1961; Porter, 1963). IgG is the most abundant class in plasma and the only that has been used for therapeutic purposes so far. The general structural features of an antibody are thus explained for the "Y-shaped" IgG molecule (see figure 1). Each chain consists of several domains that share a similar structure, the immunoglobulin fold (Williams and Barclay, 1988). The light chain has an amino-terminal variable domain V_L and a carboxy-terminal (C-terminal) constant domain C_L. The heavy chain of an IgG is composed of a variable domain V_H at the N-terminus followed by three C-terminal constant domains C_H1, C_H2 and C_H3. Antibodies are flexible due to an unstructured hinge region between C_H1 and C_H2. Each arm of the "Y-shaped" immunoglobulin is composed of a light chain and V_H-C_H1 of the heavy chain. Antibodies can be proteolytically cleaved by the protease papain in the hinge region, which results in three functional fragments (Porter, 1950). The fragments that bind the antigen are termed Fabs (fragments antigen binding) and are composed of a light chain and V_H-C_H1 of a heavy chain. Each immunoglobulin monomer can bind with both Fabs to a multivalent antigen. The resulting combined synergistic strength of binding affinities is called avidity. It is especially pronounced for polymeric immunoglobulins such as IgA and IgM (see below). The two chains of Fabs are connected by an intermolecular C-terminal disulfide bond. The remaining carboxy-terminal parts of the heavy chain comprise the Fc (fragment crystallizable) portion that is responsible for immunological effector functions. The two chains of the Fc fragment are covalently linked via disulfide bonds in the hinge

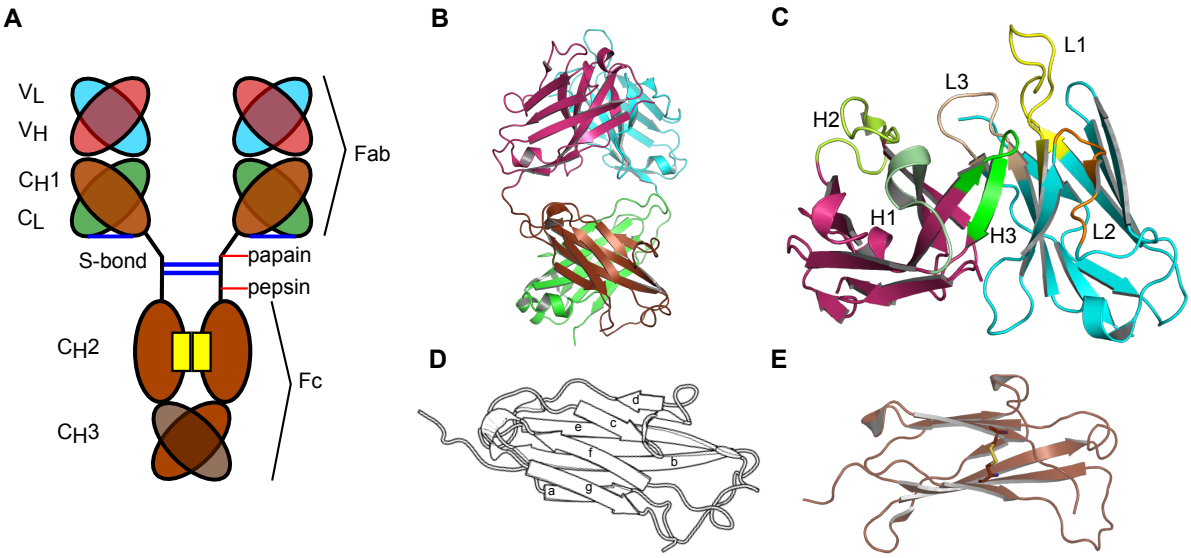


Figure 1: Structural Features of Antibodies. (A) Schematic structure of an IgG. Individual immunoglobulin domains are depicted as ellipses. Intermolecular disulfide bonds are depicted as blue lines, glycosylation of the Fc fragments is indicated by yellow squares. The cleavage sites of the hinge region for papain and pepsin are shown. **(B)** Cartoon representation of an exemplary Fab (PDB ID 3O6K), coloring comparable to (A). **(C)** The Fv fragment with the CDRs named L1 - H3. The CDRs constitute the paratope that binds to a specific region at the antigen, the epitope. **(D)** The immunoglobulin fold of domain C_H1. The β-strands are named from a-g, with a, b and e forming one, c, d, f and g composing the second β-sheet. **(E)** Each immunoglobulin domain in antibodies is stabilized by an intramolecular disulfide that connects the two β-sheets.

region and glycosylated at the C_H2 domains. When pepsin, instead of papain, is used for proteolytic digestion, the antibody is cleaved in a way that the two Fabs are held together by the intermolecular disulfide bonds of the hinge region. The bivalent fragment is called Fab'$_2$ (Nisonoff et al., 1964).

The immunoglobulin fold is a β-sandwich composed of two β-pleated sheets. In constant domains, three β-strands (five in variable domains) form one sheet and four strands form a second sheet. The loops connecting the β-strands vary in length and sequence depending on the isotypes. Although the two sheets are not connected by continuous hydrogen bonds (H-bonds), the Ig fold can also be regarded as a greek-key β-barrel. Both sheets are closely packed through hydrophobic interactions and connected via an intramolecular disulfide bond which significantly stabilizes the Ig fold (Goto et al., 1979). In humans, there are two different isotypes for the light chains, called κ and λ, that are encoded on different chromosomes but there is no functional distinction. The ratio of both chains varies across species, with the κ chain constituting 95% and 60% of light chains in mice and humans, respectively (Janeway et al., 2001). The isotype of the heavy chain determines the effector function and thus the functional class of an antibody. There are five isotypes γ, μ, α, δ and ε for the

heavy chain resulting in the classes IgG, IgM, IgA, IgD and IgE, each with different functions.

IgG IgG is the predominant isotype in human serum (80%) and has the longest half-life of all isotypes in plasma. Depending on the constant region of the heavy chain, especially C_H1 and C_H3, subclasses were identified and named IgG1 - IgG4, according to their concentration in western European populations, with IgG1 being the most abundant. IgG1 is also the most widely used subclass for therapeutic antibodies (see section 1.1.3). The subclasses differ by the number of disulfide bonds in the hinge region, flexibility and functional affinity for complement or the three FcγR (Fc-γ receptor) classes (Schroeder Jr and Cavacini, 2010). The type of antigen neutralized is affected by the subclass. IgG2 and IgG4 are associated with lipopolysaccharide antigens originating from bacteria, whereas IgG1 and IgG3 are induced by protein antigens. IgG3 has also been shown to neutralize HIV more effectively than any other subclass (Scharf et al., 2001).

IgA The predominant class of antibody in human secretions like saliva, tears, mucus and breast milk is secretory IgA. It composes almost 50% of protein in the colostrum (Woof and Mestecky, 2005). IgA is monomeric in serum but converted to dimeric IgA prior to mucosal secretion. The monomers are joined by the J chain via their C_H3 domains and by an additional secretory component that binds to one of the C_H2 domains. The secretory component anchors IgA in the mucus and protects it from proteolytic cleavage. IgA is thus the most stable immunoglobulin in secretions.

IgD IgD is the second-least abundant antibody in human serum composing less than 1% of immunoglobulins. The function of serum IgD is unclear as it does not display effector functions and is sensitive to proteolysis in its hinge region. It is also expressed as a membrane bound form on mature naive B cells when they leave the bone marrow. It acts, together with IgM, in B cell receptor signaling and has been proposed to regulate B cell fate, too. During B cell maturation, IgD gets replaced by other isotypes (Geisberger et al., 2006).

IgE Like IgM, IgE has an additional C_H4 domain. It responds for example to pathogenic worm infections by recruiting eosinophils. With less than 0.003% of all immunoglobulins, IgE is the least abundant antibody idiotype in human serum. Nonetheless it is the subtype associated with the majority of allergic reactions. IgE binds with high affinity to the IgE receptor FcεRI of basophils, mast cells, Langerhans cells and

eosinophils, thereby causing acute inflammatory response. It also promotes upregulation of FcεRI in these cells. Later exposure of the same antigen to IgE located on mast cells and basophils results in degranulation and release of histamines and other inflammatory mediator molecules. This is the cause for hypersensitivity and allergic reactions (Sutton and Gould, 1993).

IgM The first antibodies expressed during immune response are always of the IgM subtype. As these early antibodies are produced from B cells that have not yet undergone somatic hypermutation, IgM displays a reduced affinity compared to matured immunoglobulins. Because of reduced affinity and selectivity, one single type of IgM can bind to a variety of antigens. Secreted IgM is composed of five monomers that are connected by intermolecular disulfide bonds. It can bind especially well to multivalent antigens and compensate for the low affinity of the monomers by avidity effects. In addition, because of its multivalency, it can effectively clump antigens and activate complement (Niles et al., 1995). Like IgE, it has an additional C_H4 domain. The pentameric secreted IgM contains a J chain that connects two of the monomers. As for IgA, this facilitates the secretion at mucosal surfaces. In its monomeric form, IgM acts as the main antigen receptor of mature B cells, together with IgD (Goding et al., 1977).

1.1.3 Antibodies and Fabs for Therapeutic Use

The high affinity and specificity of antibodies allows to some extend the realization of Paul Ehrlich's concept of a "magic bullet" that specifically targets disease-causing molecules. Especially the modular structure of antibodies and their high stability makes them promising protein therapeutics. The introduction of a procedure for the cultivation of hybridoma cells hybrid cell lines (that originate from a fusion of antibody-producing B-cells and immortal myeloma cells) allowed for the first large-scale production of monoclonal antibodies (mAbs) (Köhler and Milstein, 1975). These antibodies were murine proteins and were thus recognized as foreign molecules when injected into patients. The ensuing elimination by the patient's immune system as well as allergic reactions and diminished effector functions restrict the therapeutic use of murine antibodies. To overcome these limitations, three major approaches exist to generate antibodies which consist primarily of human sequences or are completely human. Consequently, the majority of mAbs that are currently used in clinic is either human or humanized (Nelson et al., 2010).

With antibody engineering it is possible to swap the antigen-binding V_L and V_H domains from a murine mAb to the constant domains of human mAbs. The resulting

chimeric mAbs are 70% human and have a human Fc portion that allows for full immunological effector functions (Neuberger et al., 1985). As chimeric mAbs still contain 30% of murine sequences, they are immunogenic in patients. A reduction of immunogenicity can be achieved by humanization, i.e. by grafting only the CDR loops from the murine mAb to a human framework (Jones et al., 1986). Humanized antibodies are 85-90% human and far less immunogenic compared to chimeric mAbs. A drawback is that antibodies often loose their affinity during humanization, but this can be restored by affinity maturation *in vitro* (e.g. by phage display). A complementary approach to generate completely human antibodies are transgenic animals. A variety of mouse strains was created where the murine immunoglobulin loci were replaced by genes encoding human antibodies (Green et al., 1994; Lonberg, 2005). Upon immunization, these mice produce human, *in vivo* matured antibodies that often need no further affinity improvement. The mAb-producing B-cells can be subcloned and cultivated with well-established hybridoma methods. Transgenic mice are now widely used to generate fully human mAbs, but when the antigen is either toxic or shares a high degree of homology to the murine ortholog this technique fails.

Display technologies are the only methods capable of generating human antibodies against lethal toxins. In contrast to the use of transgenic mice, they also allow for the generation of antibodies against antigens that are highly homologous to murine proteins (Chames et al., 2009). Based on the successful application of antibody engineering technologies it became possible to clone the entire repertoire of human antibody genes. The selection of these completely human antibodies based on the affinity to the antigen or other criteria (like e.g. thermostability) is achieved with display technologies. While yeast and ribosome display are rarely used, phage display (McCafferty et al., 1990) has become the method of choice for the *in vitro* generation of human mAbs. Antibody fragments are cloned as a fusion with phage coat proteins. The resulting libraries consist of up to billions of phages each displaying a different antibody fragment. The physical link between a phenotype (e.g. binding affinity) and genotype (gene encoding the antibody) allows for the selection of phages bearing the desired genes via immunopurification. After selection of antibodies, the affinity can be increased further. Error-prone PCR that introduces mutations to the selected sequences and several rounds of subsequent phage panning are often used for this *in vitro* affinity maturation. A therapeutic antibody that was generated with phage display is adalimumab (Humira; Abbott Laboratories). It binds to tumor necrosis factor (TNF) and is now widely used for treatment of rheumatoid arthritis (Carter, 2006).

Compared to small molecule drugs, mAbs have a variety of both, advantages and shortcomings. The methods described above have paved the way for the generation of monoclonal antibodies against any desired antigen. Besides for proteins and

saccharides, this is also possible for small molecules, that are not able to evoke an immune reaction by themselves. But these haptens can be coupled to a carrier molecule (usually a protein of foreign origin) and the hapten-carrier adduct can be used for immunization. Alternatively, phage display can be used to generate mAbs against haptens.

Antibody therapeutics are often safe and well tolerated and have a high rate of success concerning clinical approval. Between 25 - 30% of antibodies that enter clinical phase I finally get approved, whereas only 11% of small molecules reach approval (Kola and Landis, 2004). Especially for targets that have no enzymatic activity or only shallow binding sites and that are thus considered "undruggable" by small molecules, antibodies offer a remedy. mAbs are usually able to bind these targets, due to a much larger interaction area. Because of the numerous atomic interactions with the antigen, antibodies usually have high affinity (with best K_Ds in the sub picomolar range reported) and selectivity. As mAbs exceed the renal clearance threshold of 70 kDa and are salvaged from lysosomal degradation by interaction with the receptor FcRn (neonatal Fc receptor), their average serum half-life of 21 days (for isotype IgG1) is usually much longer than for small molecule drugs (Ober et al., 2004). In addition of merely blocking enzymatic activities or protein interactions, mAbs can also trigger immunological effector functions. Antibody dependent cellular cytotoxicity (ADCC) and complement dependent cytotoxicity (CDC) can result in increased efficacy. When mAb is bound to an antigen on the surface of a cell, via ADCC natural killer cells as well as macrophages, neutrophils and dendritic cells are recruited which results in apoptosis and phagocytosis of the marked cell, respectively. Via CDC the complement system is activated which may lead to the formation of pores in the membrane of marked cells, although this process has not been demonstrated for therapeutic mAbs so far *in vivo* (Chames et al., 2009).

One of the disadvantages of antibodies compared to small molecule drugs is the limitation to targets that are present in the plasma or on the surface of cells or pathogens. Intracellular targets are not accessible for mAbs, as their size prevents membrane passage. In addition, as immunoglobulins are proteins, they are not orally available but must be administered intramuscularily, subcutaneously or intravenously. The lack of convenience usually translates into reduced patient's compliance. Together with high manufacturing costs and large total doses (Chames et al., 2009), this limits the use of mAbs to severe diseases. Furthermore, mAbs that have an affinity that is too high might not be uniformly distributed at the target site. This has been shown for a panel of antibodies against the tumor antigen ERBB2. The localization of mAbs at tumor cells was inversely correlated with their affinity, with the best binding mAbs located only at the perivascular region of the tumor (Adams et al.,

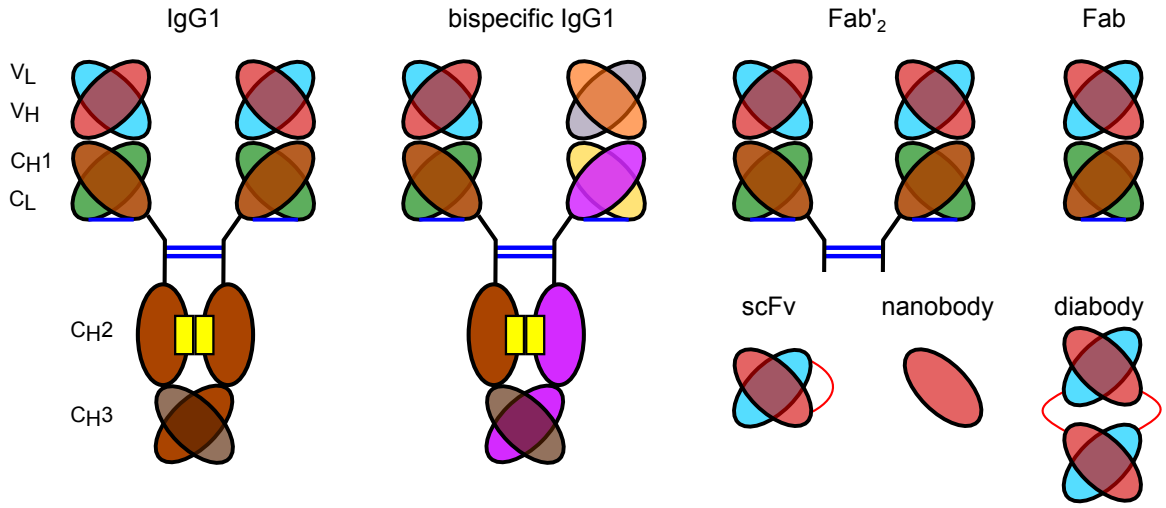


Figure 2: Formats of Therapeutic Proteins Based on Antibodies. An exemplary collection of some of the antibody formats that already entered clinical studies is shown. IgG1: Immunoglobulins of the isotype G1 still constitute the majority of all therapeutic antibodies. bispecific IgG1: bispecific antibodies promise to simultaneously bind two separate targets. This may be beneficial e.g. for recruiting or activating effector cells. Fab'$_2$: Bivalent Fab fragments can be generated by digestion of full-length antibodies with pepsin. While they can no longer trigger effector functions, they still have a high affinity because of avidity effects. Fab: Antigen binding fragments have only about on third the size of an IgG. They are hence cleared much faster by renal elimination. Fabs are monovalent binders and thus not capable of inducing receptor activation by cross linking. scFv: single chain Fvs are variable domains of light and heavy chain covalently connected by a peptide linker. Because of their small size they display improved tissue penetration especially in tumor therapy. Nanobody: The variable domains of cameloid or shark heavy chain antibodies are marketed as nanobodies. Compared to other fragments, nanobodies have increased gastrointestinal stability and are to some extend orally available. With a size of only 12 kDa, they are also expected to be able to bind to epitopes not accessible for larger fragments. Diabody: two scFvs with shortened linkers are able to form diabodies. They have the same molecular size and pharmacological features as Fabs, while providing bivalency and hence improved affinity due to avidity (adapted and modified from Carter, 2006).

2001).Despite these disadvantages, human monoclonal antibodies are a fast growing class of drugs, with full-length IgG1 representing the most widely used isotype.

Alternative formats or antibody fragments can be employed, if altered physiochemical or pharmacological features are desired. Although there is growing interest in single chain variable fragments (scFv), diabodies, nanobodies, bispecific mAbs, Fab'$_2$and other immunodomain proteins (see figure 2), there are currently only 3 Fabs and one scFv approved by the FDA. As Fabs lack the Fc portion, they are not capable of conferring immunological effector functions like ADCC or CDC. They hence can act only by direct binding to either a ligand or it's receptor. In addition, Fabs are monovalent compared to bivalent full-length IgG1 or Fab'$_2$ fragments. This can be a significant advantage in cases when ligand binding to a receptor should be blocked while at the same time cross linking and potential activation of the receptor needs

to be prevented (Stuttle et al., 1991). With a molecular weight of ~48 kDa, Fabs are mainly eliminated via renal clearance and have an average half life of less than one day (Roskos et al., 2004). Although this is a drawback in oncology therapy as it impedes efficient tumor uptake, there are applications where fast clearance is desired, like e.g. in the neutralization of toxins or drugs. Fabs are not glycosylated which permits for the use of prokaryotic expression hosts and reduced production costs.

1.2 Dabigatran Antidote

1.2.1 The Coagulation Cascade

Coagulation describes the formation of blood clots (thrombi) and is also called thrombogenesis. Thromboembolic disorders arise from thrombi obstructing blood flow in important tissues and organs like heart and brain. Thromboembolic events are triggered by an excessive activation of the coagulation system (see figure 3) which consists of a cellular and a protein-dependant component. All of these components exist under physiologic conditions in an inactive form. The proenzymes and procofactors are proteolytically cleaved to their active forms (Mann and Lorand, 1993). The coagulation factors are generally serine proteases and are sequentially numbered according to their discovery. Factors V and VIII are glycoproteins that act as cofactors upon activation, FXIII is a transglutaminase. Immediately after injury of blood vessel walls, primary hemostasis begins. Proteins like collagen or tissue factor that are isolated from endothelium in intact vessels are exposed to circulating blood. This initiates changes to platelets which then form a plug at the site of injury. Simultaneously, proteins of the coagulation system, called clotting factors, are activated in secondary hemostasis which finally results in the conversion of fibrinogen to fibrin. Fibrin rapidly polymerizes into fibrin strands that, upon cross-linking, form the fibrin clot which ultimately stabilizes the platelet plug and impedes blood loss.

Upon lesions of the endothelium, platelets specifically bind to collagen with glycoprotein surface receptors. This interaction is stabilized by von Willebrand factor (vWf) that is released by platelets and endothelial cells (Sadler, 1998). Adhesion to collagen fibrils activates the platelets. Activated platelets release ADP, serotonin, platelet-activating factor (PAF), vWf, platelet factor 4, and thromboxane A2 (TXA2), which, in turn, activate additional platelets. Activated platelets change shape from spherical to stellate, and fibrinogen cross-links aid in aggregation of adjacent platelets to form an initial platelet plug.

Secondary hemostasis comprises the coagulation cascade that consists of two pathways. The intrinsic pathway is initiated by the formation of a complex of collagen, kininogen, prekallikrein and Factor XII (FXII). The ensuing sequential activation

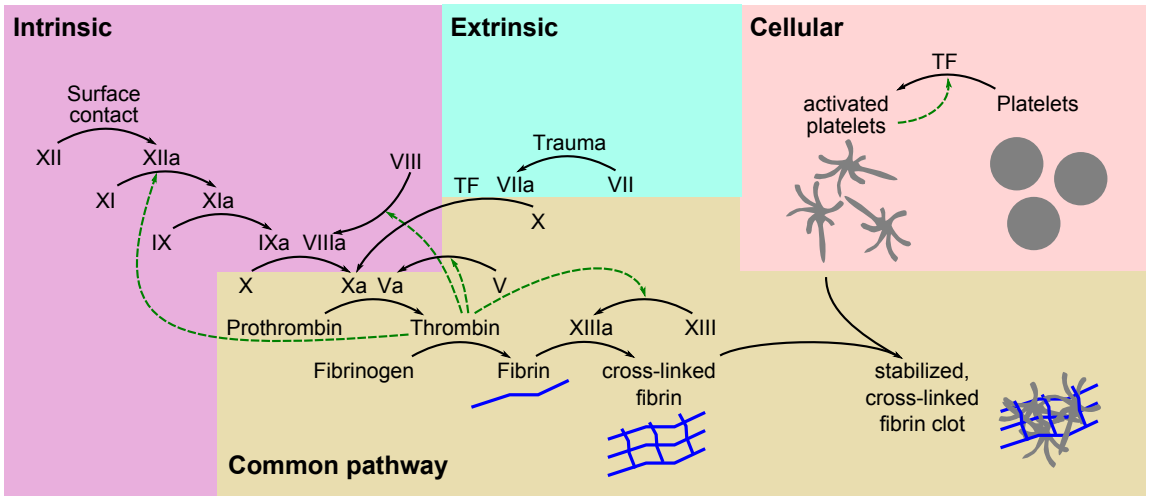


Figure 3: Coagulation Pathways. The intrinsic pathway is initiated by activation of factor XII upon contact with surface collagen. The ensuing sequential activation of downstream factors results in the activation of FX. The extrinsic pathway is activated by exposure of plasma to tissue factor (TF) upon vessel lesion. TF interacts with FVIIa and activates FX. Conversion of FX to FXa marks the first step of the final common pathway. Thrombin has a central role in coagulation as it catalyses several proteolytic conversions upstream (dotted green arrows) which result in a positive feedback loop and the thrombin burst. Thrombin converts fibrinogen to fibrin and activates factor XIII which cross-links fibrin. TF also activates platelets that change their shape to stellate and release factors that activate additional platelets. Activated platelets and cross-linked fibrin interact and by capturing erythrocytes form the final blood clot.

of FXIII, FXI and FIX results in the activation of FX and prothrombin in the common final pathway. The intrinsic pathway has only a minor role in coagulation, as patients with deficiencies of FXII, kininogen or prekallikrein do not display bleeding disorders. Instead, the intrinsic pathway seems to be linked with inflammation reactions (Chu, 2010). More important for coagulation is the extrinsic pathway, also called tissue factor pathway that generates a massive activation of thrombin, the “thrombin burst”. After lesion of a blood vessel, FVII comes into contact with with tissue factor (TF). The complex of TF and activated factor VII (FVIIa) proteolytically cleaves and thereby activates FIX and FX (Mann et al., 1998). FXa and the cofactor Va assemble to the prothrombinase complex that converts prothrombin to thrombin. The activation of FX is almost immediately inhibited by tissue factor pathway inhibitor. Thrombin initiates a positive feedback by activating factors V, and VIII which in turn activates FXI (Gailani et al., 1991) and FIX and releases FVII from the complex with vWf. FVIIIa and FIXa form the tenase complex that activates FX . The final common pathway of coagulation is initiated by the conversion of soluble fibrinogen to fibrin by thrombin (Bailey et al., 1951). By activating factor XIII, thrombin supports the covalent cross linking of fibrin polymers (Davie, 1995). Together with the thrombin-induced stimulation of platelets, this gives rise to an insoluble clot matrix. The clot

can obstruct blood flow and cause thromboembolism.

1.2.2 Anticoagulants and Antidotes

Thromboembolic disorders such as myocardial infarction, stroke and venous thromboembolism are the most common cause of mortality and morbidity in Western societies (Mathers and Loncar, 2006). The most widely used agents for antithrombotic therapy are indirect thrombin inhibitors (such as unfractionated heparin, low molecular weight heparins and fondaparinux) and vitamin K antagonists (VKAs) such as warfarin in the USA or phenprocoumon in Germany. All of these agents have certain limitations. Indirect thrombin inhibitors are mainly used when a rapid anticoagulant effect is required (Becattini et al., 2006; Hirsh et al., 2001b). The fact that these agents must be applied parenteraly provides a major obstacle for the subacute and chronic treatment. Oral available VKAs such as warfarin are primarily used for long-term out-of-hospital treatment (Hirsh et al., 2001a), although their use is cumbersome as they display multiple interactions with other drugs and food (Ansell et al., 2004). In addition, patients using VKAs need to be regularly dose-monitored because these agents have a narrow therapeutic window between the desired hemostatic and an adverse hemorrhagic effect (Jones et al., 2005). Anticoagulation therapy is always associated with an elevated risk of bleeding. This can be caused by injuries, emergency medical procedures, dosing errors or haemorrhagic tendencies. When using vitamin K antagonists, the anticoagulant effect can be reversed with vitamin K, fresh frozen plasma, prothrombin complex concentrate or recombinant factor VII (Schulman and Bijsterveld, 2007). Each of these agents has certain shortcomings including prothrombotic effects, incomplete restoration of coagulation, the danger of transfusion transmitted infections or an unclear dose-response (Fiore et al., 2001). For heparins, protamine sulfate is used as a reversal agent, but bears the risk of histamine release with ensuing hypotension and bronchoconstriction (Porsche and Brenner, 1999). As a consequence of dosing problems and bleeding risks, many patients in need for anticoagulation therapy either are disregarded or don't receive adequate therapy (Darkow et al., 2005; Waldo et al., 2005). Dabigatran etexilate is a novel, potent, non-peptidic direct thrombin inhibitor (DTI) (Hauel et al., 2002). The orally administered double prodrug is hydrolized in vivo by esterases into the active form, dabigatran (Blech et al., 2008) (see figure 4). It binds reversibly to the active site of both free and clot-bound thrombin, thus effectively preventing formation of blood clots (Di Nisio et al., 2005). Currently, dabigatran etexilate has been approved in over 60 countries for prophylaxis of thromboembolism in patients undergoing knee or hip replacements and for prevention of stroke in patients with atrial fibrillation (AF).

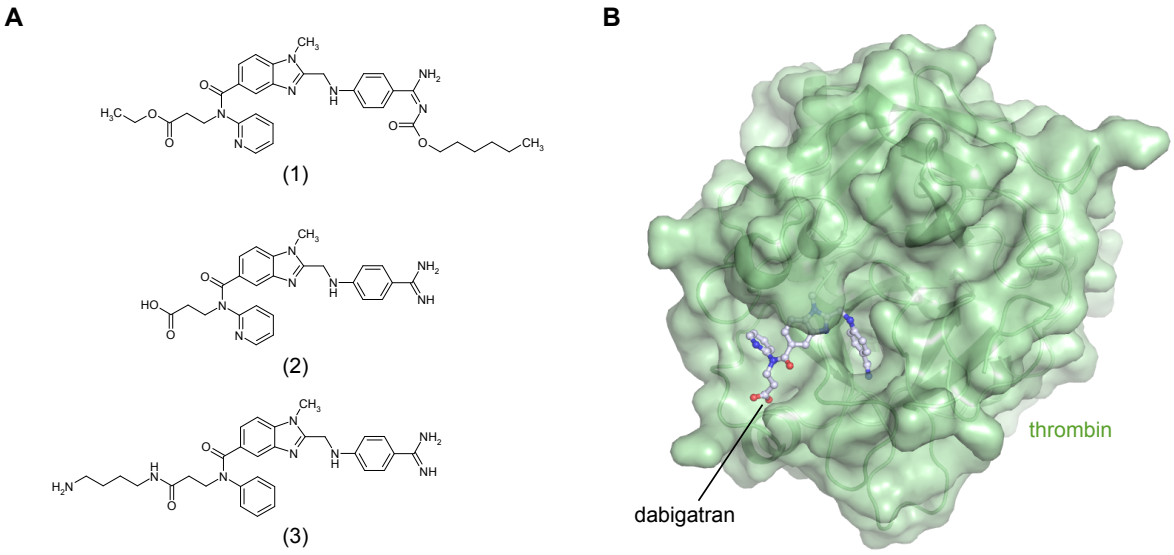


Figure 4: Structure and Mode of Action of Dabigatran. (A) Dabigatran etexilate (1) is an orally available double prodrug. Upon administration, it is hydrolyzed by esterases into the active form, dabigatran (2). For the generation of antibodies, a derivative of dabigatran (3) was coupled to a carrier protein via its amine linker and used for immunization of mice. **(B)** Dabigatran is a direct thrombin inhibitor. It binds to the active site of thrombin (green), and thereby blocks its catalytic activity.

Dabigatran treatment in patients with AF reduced the stroke risk with a better safety profile compared to warfarin (Connolly et al., 2009). In all indications, fixed-dose regimen of dabigatran has provided effective anticoagulation with a favorable bleeding profile. Compared to vitamin K antagonists, dabigatran does not require regular monitoring or dose adjustment and displays fewer adverse effects and interaction with food and other drugs (Stangier, 2008).

1.2.3 An Antidote for Dabigatran

Despite these improvements in treatment with dabigatran, anticoagulation therapy is always associated with increased risk of bleeding events. For the long-established VKAs, like marcumar and phenprocoumon, vitamin K is often used to reverse the blood-thinning effects. The emerging novel oral anticoagulants like the DTIs melagatran and dabigatran and the factor Xa inhibitors rivaroxaban and apixaban have a shorter half life and therefore shorter duration of anticoagulant effect compared to VKAs (Khoo et al., 2009). Although the use of activated charcoal or haemodialysis is recommended for the elimination of dabigatran (Van Ryn et al., 2010), there is at the moment no specific reversal agent for dabigatran or any of the other novel oral anticoagulants. Due to this unmet medical need, it was decided to develop an antidote

to reverse the anticoagulant effect of dabigatran in emergency cases. As the affinity of dabigatran for thrombin is quite high (K_D = 700 pM) (Wienen et al., 2007), a potential antidote must bind with even higher affinity to effectively inhibit interaction of dabigatran with thrombin.

Antibodies are long known to bind their antigens with extremely high affinities that can be located in the sub-picomolar range (Rathanaswami et al., 2005). Usually small molecules such as drugs cannot elicit an immune response, although they are foreign. This on the one hand is quite beneficial, as drugs would otherwise be neutralized by the immune system. On the other hand, it is not possible to generate antibodies against small molecules by simply injecting them into host organisms. Besides the display technologies mentioned in 1.1.3 on page 5 there is the hapten approach to circumvent these limitations. Following the ideas of Karl Landsteiner, a hapten is a small molecule that can elicit an immune response only when attached to a large carrier (often a foreign protein). The carrier may even originate from the organism that is used for immunization (Landsteiner, 1990). The use of hapten-carrier conjugates for immunization has been proved successful for more than 1000 different small molecules (Singh et al., 2006).

To reverse the effects of dabigatran it was decided to generate a Fab molecule, that has some pharmacological features preferred over full-length antibodies. It does not elicit undesired immunological effector functions due to the lack of the Fc portion, but acts solely through binding and neutralization of dabigatran. In addition, Fabs are eliminated very fast with elimination half-lifes of less than a day due to renal clearance. A Fab with sufficiently high affinity should thus be able to bind dabigatran and reverse thrombin inhibition. The resulting Fab:dabigatran complex should be eliminated from the body in a short period of time. There are successful examples of antibody fragments that are used as antidotes for small molecule drugs or toxins. Especially for the treatment of drug abuse-related overdoses several antibodies have been documented. Tricyclic antidepressants (TCAs) are the most common cause of death from intentional drug overdose because of hypotension and ventricular arrhythmia. A Fab' fragment that binds to the TCA desipramine was shown to rapidly reduce the cardiovascular toxicity associated with overdose in rats (Brunn et al., 1992). For the treatment of acute Cocaine toxicity the generation of an antidote in the Fab'$_2$ and Fab formats was recently reported (Treweek and Janda, 2012). The engineered fragment showed nearly complete prevention of lethality even after the appearance of acute behavioral signs of cocaine toxicity in pretreated mice. Preclinical testing of a scFv fragment against methamphetamine (also known as “meth” or “crystal”) was published in 2008 (Peterson et al., 2008). In rats the fragment was able to redistribute methamphetamine into serum but clearance of methamphetamine was

not affected as the scFv was degraded in urine. There are currently two Fabs that are used as antidotes for colchicine and digoxin in human patients. Colchicine overdoses are associated with a high mortality rate due to hemodynamic collapse. In an exemplary report a Fab was successfully used to rescue a patient that had ingested 60 mg of colchicine in a suicide attempt (Baud et al., 1995). The most widely used immunoglobulin-based antidote is a Fab against the digitalis toxin digoxin (Antman et al., 1990). The "Digoxin Immune Fab" is isolated from immunized sheep and marketed as Digibind (GlaxoSmithKline) and DigiFab (BTG). It is now the method of choice to treat severe digoxin intoxication.

1.3 PCSK9 - A Target for Dyslipidemia

1.3.1 Biological Function of PCSK9

Cholesterol is a central molecule in a variety of biological processes. It is an essential stabilizing constituent of cell membranes in mammals and a precursor for many different biomolecules like bile salts, steroid hormones and fat-soluble vitamins (e.g. vitamin D). It is formed predominantly in the liver and constitutes the majority of sterol synthesized in vertebrates. Because of its importance, the maintenance of cholesterol levels and transport throughout the body are highly regulated processes. One central step is the receptor mediated uptake of low-density-lipoprotein (LDL) cholesterol. Hypercholesterolemia describes the state of elevated levels of LDL cholesterol and is closely linked to coronary heart disease, the major cause of morbidity and death in the world (Roger et al., 2011). LDL is cleared from the blood by LDL receptor protein (LDLR) on the cell surface of hepatocytes (Brown and Goldstein, 1986). The complex of LDL and LDLR is internalized by clathrin-mediated endocytosis (see figure 5 on the next page). After migration to the endosome and subsequent acidification of the pH, LDL is released from the complex and gets degraded in the lysosome (Jeon and Blacklow, 2005), whereas LDLR is recycled and returned to the cell surface (Davis et al., 1987). Loss of function mutations in LDLR or apolipoprotein B, the component of LDL that binds to LDLR, result in familial hypercholesterolemia and coronary heart disease (Varret et al., 2008). In 2003, PCSK9 (proprotein convertase subtilisin/kexin type 9), also referred to as NARC-1 (neural apoptosis-regulated convertase 1), was discovered as the ninth member of the proprotein convertase family (Seidah et al., 2003). Proprotein convertases regulate many cellular processes through proteolysis of a variety of proteins into their active or inactive forms (Seidah et al., 2008). Gain-of-function mutations in the PCSK9 gene were identified to be linked to autosomal dominant hypercholesterolemia (ADH), and therefore as the third cause of familial hypercholesterolemia, besides mutations in LDLR and ApoB

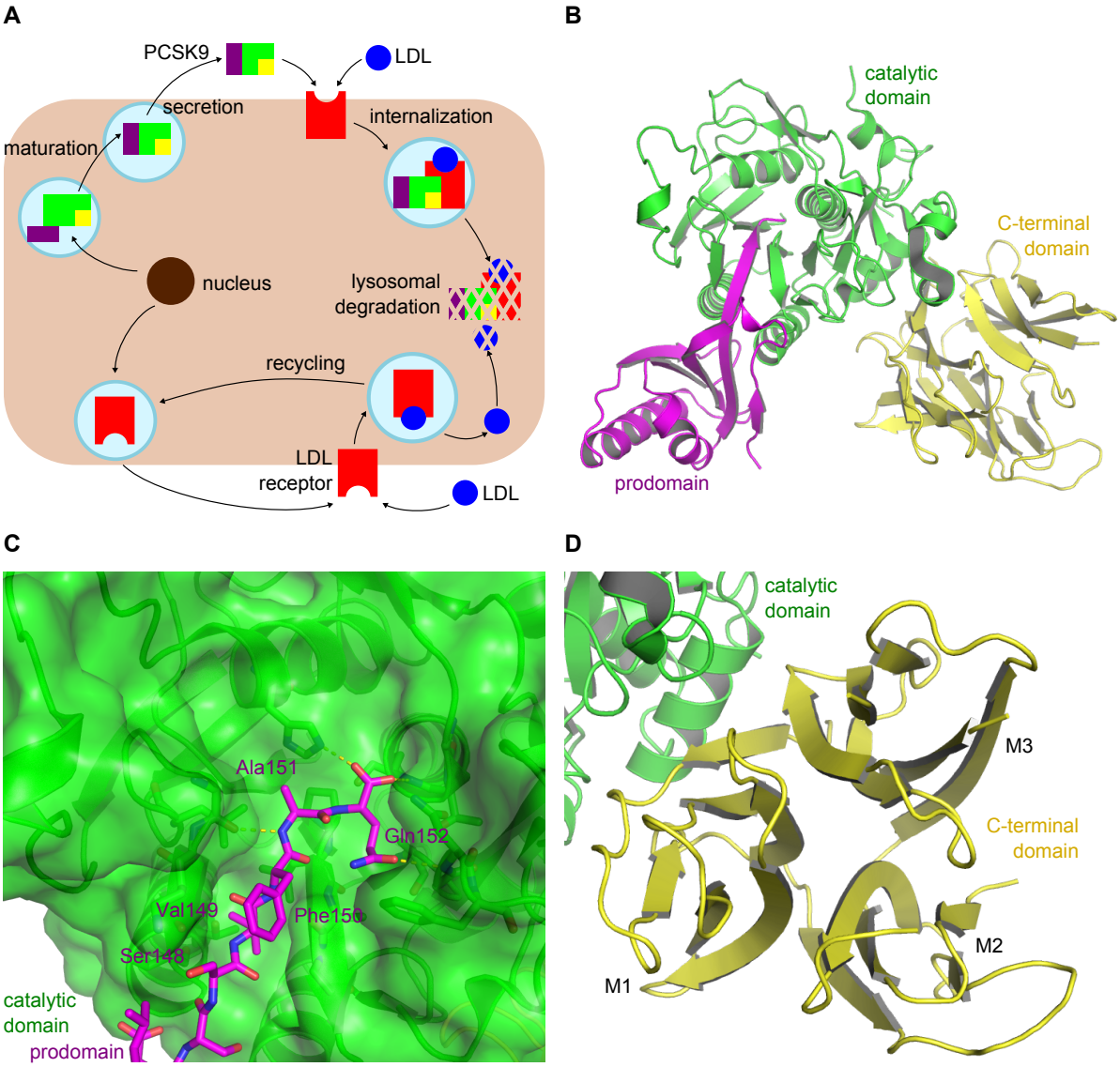


Figure 5: Function and Structure of PCSK9. **(A)** Proprotein convertase subtilisin/kexin type 9 is cleaved after the prodomain during maturation. Cleavage is necessary for secretion from hepatocytes and subsequent binding to LDL receptor (LDLR). The complex of PCSK9, LDLR and LDL is degraded via the lysosomal pathway while LDLR is recycled in the absence of PCSK9. **(B)** Matured PCSK9 is composed of 3 distinct domains: the N-terminal prodomain, the catalytic domain and the cysteine- and histinde-rich C-terminal domain. **(C)** The C-terminal part of the prodomain (purple) blocks the active site of the catalytic domain (indicated as sticks). **(D)** The C-terminal domain comprises a previously unknown fold. It consists of three very equal modules that each display 6 antiparallel strands in a truncated jelly roll motif-like fold, but have significant deviations in amino acid sequence. The modules are named from M1- M3, with M3 comprising the C-terminus.

genes (Abifadel et al., 2003). Subsequent studies demonstrated that PCSK9 overexpression in mice resulted in reduced LDLR expression in the liver and increased levels of serum cholesterol (Benjannet et al., 2004; Maxwell and Breslow, 2004; Park et al., 2004). Individuals with loss-of-function mutations in PCSK9 display a reduced risk of coronary heart disease (Cohen et al., 2006). The concentration of PCSK9 circulating in plasma is 2-4 nM (Dubuc et al., 2010; Alborn et al., 2007). Functional studies showed that PCSK9 does not degrade LDLR but binds to it on the surface of hepatocytes. In the presence of the endocytic adaptor protein ARH (autosomal recessive hypercholesterolemia) that binds LDLR in the cytosol, the whole PCSK9-bound receptor complex is internalized and undergoes lysosomal degradation (Horton et al., 2007; Lagace et al., 2006). The main interaction is formed between PCSK9 and the epidermal growth-factor-like repeat A (EGF-A) domain of LDLR (Kwon et al., 2008; Zhang et al., 2007). In 2011, Surdo et al. reported the crystal structure of PCSK9 in complex with full-length LDLR at neutral pH that provides an explanation for the mode of action of PCSK9 (Surdo et al., 2011). They identified an additional interaction between the prodomain of PCSK9 and the LDLR β-propeller domain, which keeps the LDLR in an extended conformation. At acidic endosomal pH, this inhibits rearrangements to a closed receptor conformation that is thought to be required for release of LDL and subsequent recycling of LDLR (Rudenko et al., 2002). At the same time the affinity of PCSK9 for EGF-A is enhanced at low pH (Bottomley et al., 2009; Kwon et al., 2008). Depending on the techniques used, the determined K_D of PCSK9 for LDLR is 170 - 840 nM at plasma pH (Cunningham et al., 2007; Piper et al., 2007; Fisher et al., 2007). In addition to this extracellular mode of action, there is also an intracellular pathway of PCSK9-induced LDLR degradation (Poirier et al., 2009). LDLR and PCSK9 interact early in the secretory pathway which also leads to lysosomal degradation (Nassoury et al., 2007). But the exact mechanism is still not understood completely. Nonetheless, inhibition of the extracellular PCSK9 pathway seems to be sufficient for therapeutic applications, as explained in section 1.3.3.

1.3.2 Structure of PCSK9

After cleavage of the signal peptide (residues 1 – 30), PCSK9 consists of 3 discrete domains, a prodomain PD, (residues 31 – 152), a catalytic domain CD (residues 153 – 451) and a cysteine- and histidine-rich C-terminal domain CTD (residues 452 – 692). PCSK9 has no classical P domain that is important for folding and regulation of activity in other proprotein convertases (Naureckiene et al., 2003; Zhou et al., 1998). The PD resembles the prodomain of subtilisin with two α helices and a four stranded antiparallel β sheet (Gallagher et al., 1995). All other proprotein conver-

tases have an additional cleavage site in their prodomains that allows for the release of the prodomain and results in active proteases. During maturation, PCSK9 undergoes autocatalytic cleavage between the prodomain and the catalytic domain in the ER (Naureckiene et al., 2003). This cleavage is necessary for subsequent secretion of the protein (Seidah et al., 2003). The proteolysis is performed by a catalytic triad consisting of Ser386, His226 and Asp186. The prodomain remains associated with the catalytic domain, blocks access to the catalytic triad and thereby self-inhibits its proteolytic activity, as seen in the crystal structures (Cunningham et al., 2007; Hampton et al., 2007; Piper et al., 2007). The C-terminal P6-P1 segment of the prodomain interacts with the binding cleft of the CD. P1' residue Ser153 is more than 25 Å distant from P1 residue Gln152, indicating major structural rearrangements following autocatalytic cleavage. This conformational change is assumed to be required for efficient secretion. As mentioned above, matured PCSK9 has no enzymatic activity but acts solely by binding to LDLR. The structure of the CD is remarkably similar to the other subtilases. It is composed of a seven-stranded parallel β-sheet flanked by α-helices on both sides. The CTD comprises three modules M1, M2 and M3 which are related by a pseudo three-fold rotation axis. The individual modules comprise a two sheet β-sandwich made of six antiparallel strands in a truncated jelly roll motif. Each β-sandwich is held together by three internal disulfide bonds. Besides the position of these cysteines the overall sequence conservation between the modules is quite low. While there are no structural homologues for the complete CTD, the individual modules resemble the fold of *Helix pomatia* agglutinin (HPA) (Sanchez et al., 2006). The functional role of the CTD is not fully understood yet, but it was shown to contribute to the reduction of LDLR on cell surface (Zhang et al., 2008). In a cellular LDL uptake assay, PCSK9 lacking the CTD displayed a 20 fold higher IC50 compared to the wild type (Ni et al., 2010) while there are no differences in affinity for LDLR at plasma pH (Bottomley et al., 2009). The CTD also bears numerous naturally occurring loss- and gain-of-function mutations. Besides this genetic evidence for the functional significance of this region, the CTD contains several flexible regions that may be involved in binding to other proteins and putative cofactors. Annexin A2 (AnxA2) was identified to bind PCSK9 via the CTD (Mayer et al., 2008). In addition, there is evidence for an interaction of the ligand-binding domains (LBD) of LDLR with the CTD of PCSK9 (Zhang et al., 2008). Yamamoto *et al.* proposed a two-step binding model: Following initial interaction of PCSK9 with the EGF-A domain and internalization of the PCSK9:LDLR complex, binding between CTD and LBDs takes place at endosomal low pH. It was suggested that the CTD gets positively charged by protonation of its numerous surface-exposed histidine residues at endosomal pH and thereby promotes interaction with negatively charged elements in the LBD (Yamamoto et al., 2011).

1.3.3 Inhibition of PCSK9 Function

Statins are currently used as the most successful drugs to treat hypercholesterolemia. They induce LDLR expression by inhibiting HMG-CoA reductase and cholesterol synthesis. With ongoing research after its discovery, PCSK9 has become a highly desirable target for the generation of novel cholesterol-lowering drugs because of several reasons (Duff et al., 2009): Mice lacking PCSK9 are viable and exhibit decreased plasma cholesterol levels (Rashid et al., 2005). Patients with loss-of-function mutations in PCSK9 genes display lowered LDL cholesterol concentrations with no adverse effects (Cariou et al., 2009). As mentioned before, PCSK9 interacts both intracellularly and extracellularly. But the key mediator of LDLR function seems to be secreted PCSK9 as demonstrated by Chan and colleagues (Chan et al., 2009). This makes PCSK9 a target perfectly druggable with antibodies.

Based on the current knowledge of PCSK9 function there are a number of different potential approaches for its inhibition: i) knock-down of PCSK9 gene expression; ii) inhibition of autocatalytic cleavage and processing of PCSK9 prior to secretion; iii) inhibition of the protein-protein interaction between PCSK9 and LDLR; iv) prevention of the PCSK9-mediated targeting of LDLR for lysosomal degradation. The inhibition of PCSK9 by gene silencing methods like siRNA or antisense oligonucleotides (ASOs) has been investigated by several companies. Although both methods were shown to be effective in partial lowering of LDL cholesterol levels (Frank-Kamenetsky et al., 2008; Gupta et al., 2010), no gene silencing approach has entered clinical development so far. The use of small molecule drugs to inhibit the interaction of PCSK9 with LDLR remains very difficult as seen for other examples of protein-protein interactions. Nevertheless this approach was successful (among other examples) in disrupting the interaction of interleukin 2 (IL2) and its receptor IL-2Rα (Tilley et al., 1997). But small molecules may be better suited for the inhibition of autocatalytic cleavage, which is a substantial requirement for secretion. During processing in the ER, proPCSK9 acts enzymatically as a serine protease and cleaves its own prodomain. The fact that the proteolytic processing takes place early in the ER, shortly after its synthesis, might limit this approach and may be one of the reasons why at the moment no small molecule candidate has entered clinical development.

The most successful inhibition of PCSK9 function has been achieved with therapeutic antibodies. In nonhuman primates, a monoclonal antibody (mAb) against PCSK9 lowered serum LDL by 80% (Chan et al., 2009). Several other groups described antibodies that showed comparable effects, thereby giving more evidence that this is a promising strategy to treat hypercholesterolemia (Liang et al., 2012; Stein et al., 2012). All of these antibodies inhibited binding of LDLR to PCSK9 by

either direct or sterical blocking access to the EGF-A binding site. A recent report indicated that partial inhibition of PCSK9 function can also be achieved with an antibody that binds to the CTD of PCSK9 (Ni et al., 2010). This antibody inhibits PCSK9 function in cell culture assays but does not interfere with binding of the EGF-A domain to PCSK9. Rather it inhibits the internalization of the PCSK9:LDLR complex. Although the detailed mechanism is not fully understood, data suggest that the CTD needs to be accessible for the ligand binding domains of LDLR or other cofactors. This finding is in good agreement with a report that annexin A2 binds the CTD and acts as endogenous inhibitor of PCSK9 function (Mayer et al., 2008). Notably, statins have been found to up-regulate PCSK9 expression and thereby reducing their own LDL-lowering effects (Dong et al., 2010). A combination of anti-PCSK9 mAbs with statins increased LDLR levels more than either treatment alone in a cellular assay (Chan et al., 2009) and was also effective in humans (Stein et al., 2012). Such a combined treatment may be beneficial for patients who fail to reach desired LDL levels by statins alone or hypercholesterolemic patients who do not tolerate high doses of statins.

2 Thesis Motivation and Summary

The adaptive immune system uses antibodies for the highly specific neutralization of foreign pathogens. With the discovery of the hybridoma technology by Köhler and Milstein in 1975 it became possible to generate monoclonal antibodies of practically unlimited amounts. Since then, antibodies and antibody fragments became important representatives of therapeutic proteins. Especially within the last 15 years, antibodies became the fastest growing class of therapeutic molecules. Compared to small molecule drugs these proteins are not orally available and limited to act extracellular. But their large interaction area allows for high affinity binding to otherwise “undrugable” targets like e.g. enzymatically inactive proteins or proteins with flat active sites. The high development and production costs often limit their application to severe diseases or emergency situations.

A potential emergency situation may occur in patients that use “blood thinning” drugs like vitamin k antagonists (e.g. warfarin or marcumar) or next generation anticoagulants like dabigatran (a thrombin inhibitor), apixaban or rivaroxaban (both factor Xa inhibitors). As blood coagulation is reduced in these patients, severe bleeding may occur in the event of an accident or emergency surgical operations. Where it is possible to reverse the effects of warfarin and marcumar with the in application of vitamin k, there is no specific antidote for the emerging new anticoagulants. One

aim of this thesis was to support the development of an antibody fragment that binds dabigatran and neutralizes its anticoagulant effects. Therefore 3 different Fabs had to be expressed in HEK 293 cells, purified from crude sample and their binding mode had to be characterized by X-ray crystallography and biophysical methods. All Fabs were successfully purified by affinity chromatography on benzamidine sepharose and their X-ray crystal structures were determined in the free form and in complex with dabigatran. The X-ray structure offers an explanation for the extreme high affinity of one of these Fabs and revealed striking structural similarities with thrombin concerning the binding of dabigatran. Two other Fabs share less structural homology with thrombin but contain a CDR:H3 loop that drastically changes its conformation between the apo and holo structure.

The second part of this thesis focuses on antibodies that bind PCSK9, a protein involved in cholesterol homeostasis. Under physiological conditions the LDL receptor interacts with LDL cholesterol and the complex gets internalized. Upon acidification, LDL cholesterol is released and gets degraded, while the LDL receptor is recycled and gets displayed on the cell surface again. PCSK9 binds to the LDL receptor and leads to its lysosomal degradation. Reduced LDL receptor levels lead to increased levels of LDL cholesterol, a state called hypercholesterolemia. This is associated with an increased risk of atherosclerosis and coronary heart disease, the major cause of death and morbidity in the western world. Antibodies against PCSK9 have been shown to successfully decrease LDL cholesterol levels on top of statin therapy. For this thesis, the mode of action and the epitope of antibodies against PCSK9 was analyzed by structural and biophysical methods. Therefore the antibodies were expressed in hybridoma, CHO and HEK cell culture and purified with protein-A affinity chromatography. The epitope for 2 antibodies was revealed by X-ray crystallography and surface plasmon resonance. Interestingly, one of these antibodies lowered LDL cholesterol in cynomolgous monkeys while not competing with the binding of PCSK9 to the LDL receptor. The mode of action based on structural and biophysical data is discussed.

3 Material

3.1 Chemicals

Unless mentioned otherwise, all chemicals were of analytical grade purity and were purchased from Fluka (Neu-Ulm, Germany), Merck (Darmstadt, Germany), Roche (Grenzach-Wyhlen, Germany), Roth (Karlsruhe, Germany), Sigma-Aldrich (Deisenhofen, Germany), Thermo-Fisher Scientific (Schwerte, Germany) and Qiagen (Hilden, Germany).

3.2 Media

Lysogeny broth (LB)
1% (w/v) Bacto Tryptone, 0.5% (w/v) Yeast Extract, 0.5% (w/v) sodium chloride, adjusted to pH 7.5 with 2M NaOH
Super Optimal broth with Catabolite repression (SOC)
Tryptone2% (w/v) Yeast extract 0.5% (w/v), 8.6 mM NaCl, 2.5 mM KCl,
20 mM $MgSO_4$, 20 mM Glucose

3.3 Buffers

All buffers were prepared with deionized water (conductivity < 0.055 μS/cm). All buffers were passed through a nylon filter with pore sizes of 0.2 μm or 0.45 μm and degassed with a vacuum exhauster.

TAE buffer
10 mM TRIS, 1 mM EDTA, 20 mM acetate, pH 7.0
Standard protein buffer
50 mM TRIS, 100 mM NaCl, pH 7.5
Ni-NTA wash buffer
50 mM TRIS, 300 mM NaCl, 10 mM imidazole, pH 8.0
Ni-NTA elution buffer
50 mM TRIS, 500 mM NaCl, 300 mM imidazole, pH 8.0
Protein-A binding buffer
PBS
Protein-A elution buffer
50 mM Glycine, pH 3.0
BA wash buffer
50 mM TRIS, 500 mM NaCl, pH 8.0

BA elution buffer
50 mM Glycine, 1 M NaCl, pH 2.1
Papain digestion buffer
20 mM Na_3PO_4, 10 mM EDTA, 20 mM Cysteine, pH 7.0
Crystallization buffer
25 mM TRIS, 10 mM NaCl, pH 7.0
HBS-T
10 mM Hepes, 150 mM NaCl, 0.05% Tween-20, pH 7.4
PCSK9-SPR buffer
50 mM TRIS, 200 mM NaCl, 2 mM $CaCl_2$, 0.01% Tween-20, pH 7.4
KinExA buffer
PBS supplemented with 0.1% BSA
ITC buffer
PBS supplemented with 1% DMSO and 0.05% HCl.
PBS
137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4

3.4 Crystallization Material

Cryoloops	CrystalCap HT™	Hampton Research
Centrifugal devices	Nanosep® 0,2μm	PALL Life Sciences
Crystallization plates	Intelli-Plate 96-2 well	Art Rob. Instruments
	Intelli-Plate 96-3 well	Art Rob. Instruments
	VDX™48 Plate (sealed)	Hampton Research
	VDXm™ Plate (unsealed)	Hampton Research
	Linbro® Plate	Hampton Research
Cryo tools	Cryotong™ (long)	Hampton Research
	Crystalwand Magnetic™	Hampton Research
	Vial Clamp™ (straight)	Hampton Research
	CryoCane 5 vial holder	Hampton Research
Siliconized circle glass cover slides	Foam Dewar	Hampton Research
	Ø 18 mm x 0.22 mm	Hampton Research
	Ø 22 mm x 0.22 mm	Hampton Research
Plain circle glass cover slides	Ø 18 mm x 0.22 mm	Hampton Research
Cover slide handling	Circle cover slide Gizmo	Hampton Research

Seed bead™	Seed bead™ kit	Hampton Research
Crystallization instrument	Micro-Tools II ™	Hampton Research
Crystallization screens	Additive Screen	Hampton Research
	Classic I	Qiagen, Jena Biosc.
	Classic II	Qiagen, Jena Biosc.
	Index	Hampton Research
	JCSG+	Qiagen, Jena Biosc.
	Low Ionic Strength (LISS)	Hampton Research
	MIDAS	Molecular Dimensions
	PACT	Qiagen, Jena Biosc.
	ProComplex	Qiagen
	Salt Rx	Hampton Research
Liquid handling	Formulator	Formulatrix
	Mosquito	TTP LabTech Ltd.
	Pheonix	Art Robbinson Instruments
	Gryphon	Art Robbinson Instruments
Plate imager	Rock Imager 100	Formulatrix
Stereomicroscopes	SZ61, SZX16, SZ-PT	Olympus
	ZLED CLS9000	Zett Optics
	M165 C KL	Leica Microsystems
Cryotubes	Nunc Cryo Tube™ Vials	Nalge Nunc International

3.5 Materials for protein expression and purification

Affinity columns/-material	Benzamidine Sepharose™ 6B	GE Healthcare Biosciences
	HisTrap ™ HiTrap™ Butyl FF 5mL	GE Healthcare Biosciences
	HiTrap™ Q HP 1mL	GE Healthcare Biosciences
	Ni-NTA Superflow agarose	Macherey-Nagel GmbH

	Papain, immob. Sepharose	Thermo Fisher Scientific
	Protein A, immob. Sepharose	Thermo Fisher Scientific
	Superdex 200 10/300 GL	GE Healthcare Biosciences
	Superdex 200 26/60	GE Healthcare Biosciences
	Superdex 75 10/30 HR	GE Healthcare Biosciences
	Superdex 75 5/150 GL	GE Healthcare Biosciences
Analytical balance		Sartorius AG, Göttingen
Cell rupture	Constant Cell Basic-Z	Constant Systems Limited
Chromatography devices	ÄKTA™ Explorer 100	GE Healthcare Biosciences
	ÄKTA™ FPLC®	Amersham Pharmacia Biotech
	ÄKTA™ Micro	GE Healthcare Biosciences
Cryotubes	Nunc Cryo Tube™ Vials	Nalge Nunc International
Desalting columns	illustra NAP™-5 columns	GE Healthcare Biosciences
	PD-10	GE Healthcare Biosciences
Dialysis devices	D-Tube™ Maxi Dialyzer	Novagen U.S.
	D-Tube™ Midi Dialyzer	Novagen U.S.
	D-Tube™ Mini Dialyzer	Novagen U.S.
Dialysis tubes	MWCO 10,000-12,000	Carl Roth GmbH & Co. KG
	MWCO 6,000-8,000	Carl Roth GmbH & Co. KG
Disperser	Titanteller 477 TT 13	Bandelin Electronic GmbH
	Ultra-Turrax® T25 basic	IKA Werk GmbH & Co.KG
Electrophoresis chambers	XCell SureLock® Mini-Cell	Invitrogen Life Technologies

Falcon® Tubes	15mL and 50mL	Becton Dickinson GmbH
Fermenter	Techfors-S	Infors HAT
Filter devices	0,2μm und 0,45μm	Sartorius AG
Heating block	Techne DRI-Block DB-2D	Labbay BV
iBlot® Western Detection Kit	anode stack	Invitrogen Corporation
	disposable sponge	Invitrogen Corporation
	filter paper	Invitrogen Corporation
	kathode stack	Invitrogen Corporation
Incubator	Jouan	Thermo Electron Corporation
Magnetic stirrer	IKA®	IKA® Werke, Staufen
Mini Cellophan Foil	DryEase™	Invitrogen Corporation
NuPAGE 4-12 % BisTris Gel	1,0 mm x 10 well	Invitrogen Corporation
	1,0 mm x 12 well	Invitrogen Corporation
pH meter	WTW Multi-Cal System	WTW GmbH
Power supply	PowerEase® 500 Power	Invitrogen Life Technologies
Pipettes	1000μL und 5000μL	Eppendorf AG, Hamburg
	2μL, 20μL, 100μL, 200μL	
Precision balance		Sartorius AG, Göttingen
Rotator	Scientific Drive STR4	Bibby Scientific
SPR	Biacore T100	GE Healthcare Biosciences
Ultrasonic	Mikrospitze MS73	Bandelin Electronic GmbH
	Sonopuls HD 3200	Bandelin Electronic GmbH
UV/VIS spektrophotometer	Beckman Coulter Du 530	Beckman Coulter GmbH
	NanoDrop® ND-1000	NanoDrop Technologies
Vortex		VWR International, Darmstadt
Water bath	Haaka B3 FISONS	Fisher Scientific, Waltham
Western blotting	iBlot® Dry blotting System	Invitrogen Life Technologies

3.6 Software

Biacore T200 Evaluation Software	v. 1.0	GE Helathcare
CCP4 suite	v. 6.2.0	CCP4
Coot	v. 0.6.2	Miramar
GraphPad Prism5	v. 5.04	GraphPad Software Inc.
PyMOL	v. 1.3.r0	DeLano Scientific LLC.
RockMaker	v.2.3.0.164	Formulatrix Inc.
Unicorn		GE Healthcare
VectorNTI	v. 10.3.1	Invitrogen

3.7 X-ray equipment

Automation for crystal screening

liquid handling/dispensing Formulator™ (Formulatrix), Mosquito® (TTP LabTech Ltd.), Pheonix and Gryphon (both from Art Robbinson Instruments, Sunnyvale, CA, USA)

Plate storage and imaging

Rock Imager 100 (Formulatrix)

Preparation

arium® pro VF (Sartorius Stedium Biotech), Centrifuge 5804R (Eppendorf AG, Hamburg, Germany), Mix Mate® (Eppendorf AG, Hamburg, Germany)

Stereomicroscopes

Olympus SZ61, Olympus SZX16, and Olympus SZ-PT (all purchased from Olympus), ZLED CLS 9000 (Zett Optics), M165 C KL 1500 (Leica Microsystems)

X-ray generator unit 1

MicroMax-007 HF High Intensity Microfocus Rotating Anode X-ray generator (Rigaku) equipped with a Saturn 944 CCD (Saturn) and Mar 345 DTB (Marresearch) Image plate detector, a VariMax HF (Rigaku) optic, a 4-axis partial χ AFC-11 goniometer

(Rigaku) for the CCD detector and a Cobra non liquid nitrogen cooling System (Oxford Cryosystems). The instruments were controlled either by the JDirector (CCD) or the Mar 345 DTB software.

X-ray generator unit 2

Incoatec Microfocus Source IμS generator (Incoatec) equipped with a Mar 345 DTB (Marresearch) Image plate detector, a VariMax HF (Rigaku) optic, and and a Cobra non liquid nitrogen cooling System (Oxford Cryosystems). The instrument was controlled by the Mar 345 DTB software.

SLS beamline X06SA-PXI

In-vacuum undulator in short straight section 6S, a LN_2 cooled fixed-exit Si(111) monochromator for sagittal and a dynamically bendable mirror for meridional focussing, a high resolution diffractometer (HRD), an Oxford Cryojet (Oxford Cryosystems) and a Pilatus 6M hybrid pixel detector (Dectris Ltd.). The instrument was controlled by SLS proprietary software.

SLS beamline X06DA-PXIII

SLS super-bending magnet with 2.9 Tesla, a vertically collimating mirror, a Bartels Monochromator with dual channel cut crystals (DCCM) in (+−+) geometry, and a toroidal mirror (M2) to vertically and horizontally focus the beam at the sample position, an AEROTECH high precision air-bearing rotation axis (ABR) and a PRIGo goniometer, an Oxford Cryojet (Oxford Cryosystems) and a PILATUS 2M-F hybrid pixel detector (Dectris Ltd.). The instrument was controlled by SLS proprietary software.

4 Methods

4.1 Molecular Biological Methods

4.1.1 Preparation of Plasmid and Linearized DNA

Plasmid DNA was isolated from *E. coli* with the NucleoSpin Plasmid Kit (Macherey-Nagel) or HiSpeed Plasmid Maxi Kit (Qiagen). The detailed procedure is given in the manual, thus only the basic principle is explained here. It follows mainly the alkaline extraction method described by Bimboim and Doly (Bimboim and Doly, 1979) and subsequent binding of plasmid DNA to silica membrane (Vogelstein and Gillespie, 1979). After centrifugation of an overnight culture of bacteria containing the plasmid with the gene of interest, the pellet is resuspended in buffer supplemented with RNase and lysozyme. Lysis is completed by adding an alkaline buffer containing SDS and NaOH. In the narrow pH range adjusted (12.0 – 12.5), linear and chromosomal DNA is denatured but not plasmid DNA. Upon addition of a neutralization buffer containing high concentrations of sodium acetate, chromosomal DNA aggregates and forms an insoluble network and appropriate pH conditions for binding of plasmid DNA to the silica membrane are adjusted. Simultaneously, the high concentration of SDS causes precipitation of protein-SDS complexes which are, together with chromosomal DNA, pelleted by centrifugation. The supernatant is then transferred onto a NucleoSpin Plasmid column and contaminants like salts and other components are removed by a washing step. Plasmid DNA is selectively eluted under low ionic strength (5 mM Tris/HCl, pH 8.5). The concentration of DNA samples was determined by specific absorbance at a wavelength of 260 nm. Purity was verified by measuring at 280 nm, the characteristic wavelength for aromatic amino acids of proteins. Samples were judged as pure and used further if the ratio of A_{260}/A_{280} was between 1.7 – 2.0.

4.1.2 DNA Restriction by Endonucleases

Restriction endonucleases are capable of cutting DNA at specific restriction sites. Type II endonucleases that cut double-stranded DNA are used most commonly for cloning purposes. These enzymes often recognize palindromic sites with a length of 4-8 nucleotides and usually cut within these sites. The hydrolysis of phosphodiesters in DNA results either in fragments with blunt or sticky ends (Roberts and Murray, 1976). Restriction digestion was used for analytical and preparative purpose. The buffers recommended by the vendors for optimal reaction conditions were used during restriction at 37 °C for 1 h. The general composition of a typical digestion reaction is

Table 1: Composition of Analytical and Preparative DNA Digestions with Endonucleases.

	analytical	preparative
DNA [µg]	10	100
10x restriction buffer [µl]	2	3
restriction enzyme 10 (U/µl) [µl]	1 each	1 each
dd H_2O	ad 20	ad 30

shown in table 1. Resulting fragments were separated according to their length by agarose gel electrophoresis and cloned into the desired pTT5 vector. For cloning of aPCSK9-Fab3 into the pTT5 vector, the restriction enzymes EcoRI and ClaI were used.

4.1.3 Agarose Gel Electrophoresis

Agarose gel electrophoresis was used for the separation of DNA fragments for both, analytical and preparative purposes. DNA is negatively charged at physiological pH of 7.0 and moves towards the anode in the electrical field applied to the gel. The fragments are separated according to their conformation and molecular weight. The electrophoretic mobility is inversely correlated to the molecular weight of the DNA (Southern, 1979). DNA bands in the gel are detected with the dye ethidium bromide, which fluoresces under ultraviolet light after intercalation (Sharp et al., 1973). Ethidium bromide-stained E-Gel precast agarose gels (Invitrogen) and ReadyAgarose mini gels (Bio-Rad) were run in TAE buffer for analytical and preparative electrophoresis, respectively. The DNA samples were mixed with DNA loading buffer and applied into the wells of the gel. Gels were run at a constant voltage of 100 V for 30 min and subsequently analyzed with an UV-visualizer. The desired fragments were cut from preparative gels with a scalpel and isolated using the NucleoSpin Extract II kit (Macherey-Nagel). It follows the method introduced by Vogelstein and Gillespie (Vogelstein and Gillespie, 1979) where the agarose matrix is destroyed and DNA binds to a silica membrane in the presence of a chaotropic salt. The mixture was then loaded onto a NucleoSpin Extract column and the DNA was isolated as described in section 4.1.1.

4.1.4 Ligation

Ligation describes the formation of a phosphodiester bond between a 5'-phosphate group and a 3'-hydroxyl group of double-stranded DNA, catalyzed by the enzyme DNA ligase. Under physiological conditions, ligases are used to repair double-strand and single-strand breaks in the DNA. For cloning purposes, these enzymes are used to join inserts and vector DNA with complementary ends. After linearization and

isolation, the vector back-bone is dephosphorylated to prevent self ligation of the vector in the next step. Rapid DNA Dephos & Ligation kit (Roche) was used for dephosphorylation of the linearized vector backbone of pTT5 and subsequent ligation with the coding regions of the light chain of Fab3H42 and the truncated heavy chain of aPCSK9-Fab3, respectively. All DNA fragments were isolated from an agarose gel and purified as described. All steps were performed according to the instructions of the manufacturer.

4.1.5 Transformation and Culture of *E. coli* Bacteria

The introduction of foreign DNA into bacteria is called transformation. During cloning transformations are performed to amplify and select plasmids. 100 μl of Library Efficiency DH5α Competent Cells (Invitrogen) were thawed on ice and 1 μl of DNA (1-10 ng) was added. The suspension was incubated on ice for 30 min and heat shocked in a water bath at 42 °C for 2 min. Following cooling of the cells on ice for 2 min, 900 μl of SOC medium (room temperature was added and the mixture was shaken at 37 °C for 1 h. 20 μl and 200 μl of the transformed cells were plated on prewarmed agar plates with antibiotics and incubated at 37 °C over night. Selected clones were used to inoculate 4 ml LB medium with appropriate antibiotics in a Round-Bottom Tube (Falcon) and incubated at 37 °C for 6 h. DNA from these cultures was isolated and transmitted to MWG OPERON for sequencing. Positive clones were used to inoculate 250 ml of LB medium in a 1 l shaking flask and incubated with agitation for 15 h at 37 °C. DNA was extracted from these larger-scale cultures for subsequent transfection of human embryonic kidney cells. For long-term storage of positive *E. coli* clones at -80 °C, a glycerol stock was created by mixing equal volumes of bacterial culture and glycerol buffer.

4.1.6 DNA Sequencing

DNA sequencing was performed by Eurofins MWG Operon based on a modification of the Sanger method (Sanger et al., 1975, 1977). The DNA to be sequenced is amplified by polymerase chain reaction (PCR) with statistically incorporated dideoxynucleotide chain terminators that carry flourescent dyes. The resulting fragments of different length are subsequently separated by electrophoresis. The chain terminators carrying the dyes are analyzed with a fluorescence reader and from the resulting chromatogram the sequence is derived.

Table 2: Composition of Transfection Solution for 100 ml Culture Volume.

	Solution A	Solution B
Heavy chain plasmid [µg]	33	-
Light chain plasmid [µg]	66	-
293fectin [µl]	-	100
Freestyle F17 [ml]	5.5	5.5

4.2 Cell Culture Methods

Antibodies and fragments thereof are complex biomolecules that need to be post-translationally modified to obtain full bioactivity. In contrast to mammalian cells, bacteria like *E. coli* are not capable of post translational modifications such as e.g. glycosylation, phosphorylation or sialylation. Mammalian cell lines derived from Chinese Hamster Ovary (CHO) and Human Embryonic Kidney (HEK) are currently the main choice for production of these proteins in research (Baldi et al., 2007). CHOs are usually used to generate stable cell lines for continuous large-scale protein production whereas HEK cells are more productive when transfected transiently (Galbraith et al., 2006; Haldankar et al., 2006).

4.2.1 Cultivation and Transient Transfection of HEK293-6E Cells

The human cell line HEK293 is widely used for transient gene expression because of its good transfectability and ability to grow in suspension to a high cell density in serum-free media (Durocher et al., 2002; Meissner et al., 2001). HEK293-6E cells (Life Technologies) were cultivated in serum-free HEK medium at 37 °C and 5% CO_2 in an incubator shaker (120 rpm) and split every other day to a density of 2.5 x 10^5 cells/ml. The cells were transfected by lipofection (Felgner et al., 1987) at an initial cell density of 8 x 10^5 cells/ml with expression plasmids coding for the light and heavy chain. The method is based on complex formation between cationic lipids and negatively charged DNA molecules. The DNA lipid complex is internalized and enters the endosomal degradation pathway from which it can potentially escape. After dissociation from the lipids, DNA is able to enter the nucleus and to induce protein expression (Luo and Saltzman, 2000). For transfection of 100 ml of cells solutions A and B were prepared and incubated at room temperature for 5 min (see table 2). Both solutions were mixed and incubated for 20 min at room temperature and subsequently added to the cells. After one day of incubation at 37 °C and 5% CO_2 in an incubator shaker (120 rpm) Tryptone N1 was added to a final concentration of 1% (w/v) to feed the cultures (Pham et al., 2005). 5 days posttransfection the culture was centrifuged (5000 g) and the supernatant stored at -20 °C for further use.

4.2.2 Cultivation of Hybridomas

Hybridomas are hybrid cell lines that allow for continuous production of monoclonal antibodies (mAbs). After repeated immunization of e.g. mice with the antigen, their spleens are removed and the antibody-producing B cells, that have a limited life span, are isolated. These are fused with immortal myeloma cells and selected for successful fusion (Köhler and Milstein, 1975), resulting in an immortal cell line continuously producing mAbs. Hybridoma cells expressing murine anti-PCSK9 (aPCSK9)-mAb1 were obtained from NBE research. These hybridomas were derived from mice immunized with human PCSK9. The cells were identified by screening for binding to PCSK9. 3 x 10^7 flash-frozen cells were thawed on ice and diluted in 10 ml hybridoma medium (room temperature) with 4% hybridoma enhancing supplement (HES), centrifuged for 5 min at 20 °C (100 g) to remove DMSO and resuspended in 10 ml of hybridoma medium (+ 4% HES). Hybridomas were cultured in Corning cell culture flasks at 47 °C and 7% CO_2. Cell density and vitality was inspected at regular intervals. When cell density exceeded 10^6 cells/ml, the cultures were split and passaged to inoculate fresh medium in a shaking flask with a density of 2 x 10^5 cells/ml. Further cultivation was performed in a Unitron incubator shaker (Infors) at 120 rpm, 37 °C and 7% CO_2 and the amount of HES was reduced with each passage. FBS was omitted for the final expression period of mAb1 where cells were incubated for 6 days while shaking. After final expression stage the cells were centrifuged for 10 min at 20 °C (5000 g) and the supernatant was stored at -20 °C for further use.

4.2.3 Expression of aDabi-mAb2 in CHOs

CHO cells are the most common host for large-scale manufacture of protein therapeutics as this cell line is capable of conferring natural post-translational modifications and a variety of amplification systems are available (Bollin et al., 2011). A stably transfected CHO cell line expressing human mAb C38 was obtained from NBE research. The CHOs were diluted to 3 x 10^5 cells/ml in CHO medium and incubated at 37 °C and 5% CO_2 in an incubator shaker (150 rpm). Every other day the cells were counted and split at a ratio of 1:4 when cell density exceeded 1 x 10^6 cells/ml. For production of mAb C38 a total of 3 l was inoculated with CHOs and incubated for 10 days at 37 °C and 5% CO_2 in an incubator shaker (150 rpm). After 2 days the CO_2 level was reduced to 2%. Each day feed was added (30 ml for each liter of starting volume) and the pH was measured and adjusted to be above pH 7 by addition of Na_2CO_3. From day 5 onwards, 4 ml of glucose (500 g/l) was added to each liter of medium every day. After 10 days of incubation the culture was centrifuged for 5 min at 5000 x g and the supernatant was stored at -20 °C for further use.

4.3 Protein Preparation

4.3.1 Expression and Purification of PCSK9

Expression in Insect Cells The gene encoding for human PCSK9 (uniprot ID Q8NBP7, residues 1-692) was codon-optimisied for expression in insect cells and synthesized at GENEART (Regensburg). PCSK9 cDNA fragments were then amplified by PCR and flanked by attB1 and attB2 sites. The Gateway system was used to perform cloning into pDONR221 and subsequently into pDEST8 (Invitrogen) or a modified pDEST-8 vector (gp67_Ub_TEV). The resulting open reading frame encodes for human PCSK9 amino acids 1-688 and a C-terminal LE-6His tag. This construct is referred to as PSCK9. The PCSK9ΔCTD construct consists of the gp67 signal sequence and the ubiquitin tag, followed by a TEV cleavage site and PCSK9 amino acids 31-545 with a C-terminal LE-6His tag. Recombinant baculoviruses were generated using the BAC-to-BAC system (Invitrogen). *Trichoplusia ni* cells were infected at a cell density of 1 mio cells/ml with recombinant virus in Insect Xpress media (Lonza) and the proteins were expressed for 72 h at 27 °C. Cells were collected by centrifugation and the medium containing PCSK9 was used for further purification.

IMAC Purification 6 - 8 l of frozen insect cell medium containing secreted PCSK9 was thawed in a water bath for 15 h at 8 °C. As PCSK9 has a C-terminal hexa histidine-tag, immobilized metal-affinity chromaography (IMAC) was used as a first purification step. IMAC is based on the interactions between a bivalent metal cation (Co^{2+}, Ni^{2+}, Cu^{2+} or Zn^{2+}) immobilized on a matrix and specific amino acid side chains (Bornhorst and Falke, 2000; Porath et al., 1975). Histidine exhibits the strongest interaction with immobilized metal ions as the free electrons of the imidazole ring are able to form coordination bonds. Histidine tags allow to purify proteins from a variety of expression systems, including insect cells (Kuusinen et al., 1995). Following specific binding, the proteins are eluted with a buffer containing free imidazole that competes with the His-tag for binding to the metal ions. Nickle-nitrilotriacetic acid (Ni-NTA) coupled to the solid resin agarose was employed as it displays high stability and can thus be reused several times (Hochuli et al., 1987). 3 ml of Ni-NTA per liter insect cell medium was added and incubated for 1 h on a flask roller. All purification steps were performed at 8 °C to prevent possible proteolytic degradation. The Ni-NTA resin was transferred to a glass column (500 ml) and separated from the medium by gravity flow. The resin was washed with 200 ml of Ni-NTA washing buffer to increase the final purity by removing proteins that bind non-specifically to the nickle ions (Schmitt et al., 1993). PCSK9 was eluted with 50 ml of elution buffer and complete protease inhibitor cocktail tablets (Roche) were added immediately.

Table 3: Parameters of SEC on a HiLoad-26/60-Superdex-200.

Column material	HiLoad-26/60-Superdex-200 prep grade
Column volume (=cv) [ml]	320
Column in diameter [mm]	26
Bed height [mm]	600
Equilibration [cv]	1.5
Flow rate [ml/min]	1
Fractionation volume [ml]	2
Detection [nm]	260, 280
Temperature [°C]	4

Size Exclusion Chromatography Size exclusion chromatography (SEC) allows for the separation of proteins according to their size (Lathe and Ruthven, 1955). The filtration medium consists of dextran covalently attached to cross-linked agarose that has pores of defined size (Porath et al., 1959). When a mixture of proteins is applied to these columns, small proteins are able to enter the pores of the stationary phase, whereas larger proteins can not leave the mobile phase. The effective volume to pass is thus smaller for larger proteins and they elute earlier. Small molecules are retained longer and elute later. Proteins above a certain size elute with the void volume (the fraction of volume of the mobile phase compared to the column volume).

Following IMAC, PCSK9 was concentrated to a volume of 5 ml using Amicon Ultra-15 centrifugal filter units (EMD Millipore) with a molecular weight cut-off (MWCO) of 50 kDa. Particles were removed by filtration with NanoSep MF 0.2 μm centrifugal filters (PALL Life Sciences) and the solution was applied on a Superdex-200 prep grade chromatography column (GE Healthcare Life Sciences) equilibrated and run in standard protein buffer. Detailed parameters of the chromatography are given in table 3. Purity and the amount of protein was determined by monitoring the absorption at 260 nm and 280 nm (see section 4.4.1) and SDS-PAGE (see section 4.4.2).

4.3.2 Expression and Purification of PCSK9ΔCTD

IMAC Purification 6 - 8 l of frozen insect cell medium containing secreted PCSK9-ΔCTD with a N-terminal ubiquitin sequence tag was thawed in a water bath for 15 h. All steps were performed at 4 °C. PCSK9ΔCTD has a C-terminal hexa histidine-tag and was purified by IMAC as described in section 4.3.1. After elution from the Ni-NTA column, no inhibitor was added but the protein solution was transferred to a dialysis tubing with MWCO of 10 kDa and dialyzed against 4 l of standard protein buffer for 4 h.

Table 4: Parameters of SEC on a HiLoad-26/60-Superdex-75.

Column material	HiLoad-26/60-Superdex-75 prep grade
Column volume (=cv) [ml]	320
Column in diameter [mm]	26
Bed height [mm]	600
Equilibration [cv]	1.5
Flow rate [ml/min]	1
Fractionation volume [ml]	2
Detection [nm]	260, 280
Temperature [°C]	4

Cleavage with TEV protease TEV protease form the nuclear Inclusion a (NIa) protein encoded by the tobacco etch virus (TEV) was overexpressed in *E. coli* and used to cleave off the ubiqutin tag from PCSK9ΔCTD (Carrington and Dougherty, 1988; Dougherty et al., 1988). TEV protease specifically and efficiently cleaves the recognition site GNLTFQ/G after glutamine. 1 mg of TEV protease were added per 100 mg of PCSK9ΔCTD and incubated for 15 h. Shaking of the protein solution was avoided to prevent aggregation as observed in previous trials. Proteolytic digestion was stopped by addition of complete protease inhibitor cocktail tablets (Roche).

Size Exclusion Chromatography The protein solution was concentrated to 5 ml using Amicon Ultra-15 centrifugal filter units (EMD Millipore) with a MWCO of 10 kDa and filtered with NanoSep MF 0.2 μm centrifugal filters (PALL Life Sciences) to remove any particles. Final polishing of PCSK9ΔCTD and removal of TEV protease and the ubiquitin tag was achieved by SEC on a HiLoad-26/60-Superdex-75 prep grade (GE Healthcare Life Sciences) equilibrated and run in standard protein buffer. Detailed parameters of the chromatography are given in table 4. Purity and the amount of protein was determined by monitoring the absorption at 260 nm and 280 nm (see section 4.4.1) and SDS-PAGE (see section 4.4.2).

4.3.3 Purification of Full-length Antibodies

Murine aPCSK9-mAb1 was expressed in hybridoma cells as described in section 4.2.2. Human aPCSK9-mAb2 and mAb3H42 were expressed in CHOs (see section 4.2.3) and HEK cells (see section 4.2.1), respectively. All antibodies were subsequently purified by protein A affinity chromatography. Protein A is a surface protein located in the cell wall of *Staphylococcus aureus* that binds a wide range of immunoglobulins (Forsgren and Sjöquist, 1966; Goudswaard et al., 2006). Protein A predominantly binds to the Fc part of antibodies via hydrophobic interactions and four H-bonds as seen in the crystal structure (Deisenhofer, 1981). A buffer with low pH is used to

Table 5: Incubation Times for Papain Cleavage.

antibody	incubation time [h]
aPCSK9-mAb1	5
aPCSK9-mAb2	20
aPCSK9-mAb2	7
aPCSK9-mAb3H42	21

elute immunoglobulins by changing the charge of the interface and partial unfolding (Arakawa et al., 2004). MabSelect Xtra (GE Healthcare Life Sciences), that has a recombinant protein A coupled to an agarose matrix, was used for purification of all antibodies. 10 ml of resin were added per liter of immunoglobulin containing medium and incubated for 1 h while shaking at 4 °C. The protein A resin was loaded on an empty glass Econo-Column (Bio-Rad Laboratories) and separated from the flow through. The beads were washed with 300 ml of PBS and the antibodies were eluted with 10 ml of protein A elution buffer. The low pH of the elution buffer was immediately neutralized by addition of 10 ml of 1 M TRIS (pH 8.0).

4.3.4 Papain Cleavage of Antibodies - Preparation of Fabs

Papain is a protease originally isolated from papaya that has been described to cleave immunoglobulins at their hinge region (Porter, 1959). Comparable to other cysteine proteases, a thiol residue in the enzyme's active site gets deprotonated by the adjacent basic amino acid histidine which allows for a nucleophilic attack of the sulfur atom on the substrate carbonyl carbon and subsequent cleavage of the amide bond. Papain cleaves a wide range of immunoglobulins, including IgG1, from different species such as humans, mice and rats (Mage, 1980; Rousseaux et al., 1983). The antibodies are cleaved into two Fabs (fragments antigen-binding) with a molecular weight of 48 kDa each and one Fc (fragment crystallizable) portion (Putnam et al., 1979).

All full-length antibodies were cleaved using papain immobilized on 6% beaded agarose (Pierce). The antibodies were concentrated to 4 mg/ml with Amicon Ultra-15 centrifugal filter units (EMD Millipore) MWCO 100 kDa and during concentration the buffer was replaced by papain digestion buffer. 125 μl of papain slurry per mg of antibody were equilibrated in papain digestion buffer as well. Antibodies and papain were mixed, aliquoted into 2 ml reaction tubes (Eppendorf) and placed in a shaker incubator at 37 °C. The optimal time for digestion is given in table 5. After complete cleavage of the full-length antibodies, the papain beads were removed by gravity flow through an empty PD-10 column (GE Healthcare Life Sciences). The beads were washed with half a column volume of PBS and subsequently discarded. Wash and

Table 6: Parameters of Affinity Chromatography on MabSelect SuRe.

Column material	MabSelect SuRe
Column volume =cv) [ml]	1
Column in diameter [mm]	0.7
Bed height [mm]	2.5
Equilibration [cv]	1.5
Injection volume [ml]	2
Gradient length [cv]	15
Post gradient volume [cv]	5
Flow rate [ml/min]	1
Fractionation volume [ml]	2
Detection [nm]	260, 280
Temperature [°C]	4

flowthrough fractions were pooled and further purified.

4.3.5 Separation of Antibody Fragments

Protein A binds to the Fc region of antibodies as described in section 4.3.3. Prepacked 1 ml HiTrap MabSelect SuRe columns on an Äkta FPLC system both (GE Healthcare Life Sciences) were used to separate Fabs from Fc fragments. The protein solution after papain cleavage (see section (4.3.4)) was concentrated to a volume of 5 ml with Amicon Ultra-15 centrifugal filter units (EMD Millipore) with a MWCO of 30 kDa and applied to the FPLC system equilibrated and run in protein A running buffer with a 5 ml sample loop. The Fc region was eluted with a linear gradient of protein A elution buffer. Detailed parameters of the chromatography are given in table 6.

4.3.6 Purification of Anti-dabigatran Antibody Fragments

Affinity Chromatography on Benzamidine Sepharose Antidotes for dabigatran were expressed as Fabs in HEK293-6E cells without affinity tag. These Fab originate from mice immunized with an analog of dabigatran coupled to an antigen via its propionic acid moiety. Thus it was expected that the Fabs primarily bind to the benzamidine moiety of dabigatran. Benzamidine is a well-characterized inhibitor for thrombin and other serine proteases. Benzamidine sepharose 6B (GE Healthcare Life Sciences) is 6% agarose with immobilized p-aminobenzamidine pABA), a close analog of benzamidine. 5 ml of the resin was washed with 20 ml of BA wash buffer and subsequently incubated with 1 l of medium containing the Fabs for 1 h at 4 °C on a roller. The mixture was transferred to an empty PD-10 column (GE Healthcare Life Sciences). and washed with 20 column volumes of BA washing buffer. Bound protein was eluted with 20 ml of BA elution buffer.

Size Exclusion Chromatography The Fabs were concentrated to a volume of 2 ml with Amicon Ultra-15 centrifugal filter units (EMD Millipore) with a MWCO 30 kDa and filtered with NanoSep MF 0.2 μm centrifugal filters (PALL Life Sciences) to remove any particles. Removal of possible contaminants and polishing of the Fabs was performed by SEC on a HiLoad-26/60-Superdex-75 prep grade (GE Healthcare Life Sciences) equilibrated and run in standard protein buffer. Detailed parameters of the chromatography are given in table 4 on page 35. Purity and the amount of protein was determined by monitoring the absorption at 260 nm and 280 nm (see section 4.4.1) and SDS-PAGE (see section 4.4.2).

4.4 Characterization of Proteins

4.4.1 Determination of Protein Concentration by Photometry

Beer-Lambert law states that there is a logarithmic dependence between the transmission T of light trough a solution and the molar extinction coefficient ϵ, the concentration c of the substance and the path length l (Beer, 1852).

$$T = \frac{I}{I_0} = 10^{-\epsilon l c} \tag{1}$$

where I_0 and I is the intensity of the transmitted and incident light, respectively. The transmission can be expressed in terms of absorbance A, which is for liquids defined as

$$A = -\log_{10}\left(\frac{I}{I_0}\right) \tag{2}$$

which implies that the absorbance is linearly related with protein concentration according to

$$A = \epsilon l c \tag{3}$$

Measurement of the absorption at a wavelength of 280 nm allows to determine the concentration of proteins in solution with a modification of the Edelhoch method. Edelhoch determined extinction coefficients for Trp and Tyr by using blocked amino acid analogs as model substances to represent the situation in proteins. N-acetyl-L-tryptophanamide and glycyl-L-tyrosylglycine were used for Trp and Tyr, respectively and the values were determined in pH 6.5, 6 M guanidium hydrochloride, 0.02 M phosphate buffer (Edelhoch, 1967). It was found that these values are valid for cal-

Table 7: Molar Masses and Extinction Coefficients of Proteins Used.

protein	molar mass [Da]	length [n_{aa}]	extinction coefficient [$M^{-1}cm^{-1}$]	absorption 0.1%
PSCK9	71698	666	56515	0.788
PCSK9ΔCTD	46910	432	35910	0.766
EGF-A	4748	44	1865	0.393
aPCSK9-mAb1	144323	1313	220765	1.530
aPCSK9-Fab1	47181	436	76485	1.621
aPCSK9-mAb2	146466	1344	207360	1.416
aPCSK9-Fab2	48194	449	67895	1.409
aPCSK9-mAb3H42	145303	1342	209420	1.441
aPCSK9-Fab3H42	46586	439	68800	1.477
aDabi-Fab1	47775	444	73395	1.536
aDabi-Fab2	47796	441	69385	1.452
aDabi-Fab3	47667	441	69385	1.456

culating the extinction coefficients of the denatured protein with good approximation and they can also be used to calculate the extinction coefficients of the native protein (Gill et al., 1989). To improve the accuracy, Pace and colleagues used 116 measured molar extinction coefficients from 80 proteins to calibrate the equation for the native protein in water (Pace et al., 2008). The molar extinction coefficient ϵ can be inferred from the weighted sum of all extinction coefficients of tryptophanes, tyrosines and cystines.

$$c = \frac{A_{280}}{\epsilon d} \tag{4}$$

$$\epsilon = (n_{Trp} \cdot 5500 + n_{Tyr} \cdot 1490 + n_{Cys} \cdot 125)\, M^{-1}cm^{-1} \tag{5}$$

The absorption at 280 nm was measured with a NanoDrop 1000 spectrophotometer (NanoDrop) to determine protein concentration. The respective molar extinction coefficients are shown in table 7.

4.4.2 SDS-PAGE

Sodium dodecylsulphate polyacrylamide gel electrophoresis (SDS-PAGE) was used to control the purity of proteins and complexes (Laemmli et al., 1970). The anionic detergent SDS denatures proteins and binds to them with a ratio of 1.4 g SDS/g protein. Thus proteins are masked by SDS resulting in micelles with a constant ratio between charge and molecular weight (Lottspeich and Zorbas, 1998). When an electric field is

applied to the gel, the negatively charged protein SDS micelles move through a polyacrylamide matrix of defined pore size. The separation of a protein mixture depends on the electrophoretic mobility that is defined by shape and size of the molecules. The migration length is proportional to the logarithm of the molecular weights of the proteins. NuPAGE Bis-Tris precast gels 12% (Life Technologies) were used and run in NuPAGE MES SDS running buffer (Life Technologies), a system that is based on the improved method from Schägger and Jagow which suppresses the possible reoxidation of reduced cysteines and allows for a higher resolution and improved band distribution (Schägger et al., 1987). Depending on the sample, a total of 5 - 50 μg of protein was mixed with 5 μl of Laemmli buffer (5x) and diluted with water to a final volume of 25 μl. Protein samples were heated for 5 min at 75 °C to prevent possible cleavage of Asp-Pro bonds (Kurien and Scofield, 2012). Samples were then loaded to the gel together with a molecular weight standard and run for 45 min at a constant voltage of 150 mV. Gels were subsequently removed from the tray and stained with 10 ml of InstantBlue (Expedeon) for 1 h on a rocking shaker and dried in DryEase Mini cellophane foil (Life Technologies) for further storage.

4.4.3 ELISA

Enzyme-linked immunosorbent assay (ELISA) was used to classify analogs of dabigatran according to their affinity for aDabi-Fab1 (Engvall et al., 1971; Van Weeman and Schuurs, 1971). In a competitive ELISA, aDabi-Fab1 was coupled to the plate and binding of peroxidase (POD)-conjugated dabigatran was competed off by increasing concentrations of the analogs. All steps were performed in PBS buffer. 100 μl of aDabi-Fab1 (1 μg/ml) was immobilized in each well of MaxiSorp 96 well plates (Nalge Nunc International) over night at 4 °C. The plates were washed 3 times with PBS supplemented with 0.5% Tween-20. The wells were blocked with 200 μl of BSA (3% w/v) over night at 4 °C and again washed 3 times. 3-fold dilution series of the compounds were prepared and incubated with POD-conjugated dabigatran (30 fM) in low-binding 96 well plates (Nalge Nunc International) for 2 h. From the mixture 100 μl was transferred to the plates coated with aDabi-Fab1 and incubated for 1 h at 4 °C, the plates were subsequently washed 3 times with PBS and 0.5% Tween-20. 100 μl of 1-Step Ultra TMB (Thermo Scientific Pierce) was used to determine POD activity according to the manufacturers instructions. 3,3´,5,5´-tetramethylbenzidine (TMB) is a chromogen that is colored blue when oxidized, typically as a result of oxygen radicals produced by the hydrolysis of hydrogen peroxide by POD (Josephy et al., 1982). With the addition of 100 μl of sulfuric acid (1 M), the reaction is stopped and the color of the substrate changes to yellow. The final concentration of oxidized TMB

was determined by it's absorption at 450 nm in an Infinite M200 plate reader (Tecan Group).

4.5 Protein Crystallography

4.5.1 Crystal Lattice and Lattice Symmetry

Crystals are defined by a periodic pattern of molecules extending in all three spatial dimensions (for the further descriptions vector and complex numbers will be written in bold characters). The repeating unit that builds up the crystal and contains all structural information is called unit cell (UC). It is defined by three basis vectors extending from a common origin in a right-handed system (**a**, **b**, **c**). Their magnitudes (a, b, c) and the three angles (α, β, γ) enclosed by the basis vectors are called the unit cell constants. The smallest unit of a crystal that can generate the complete crystal structure by means of its symmetry operations is called the asymmetric unit (ASU) of the unit cell (Blow, 2003). The complete group of symmetry operations that generate the three dimensional lattice and defines the symmetry of the crystal is called space group. There exist 230 three dimensional general space groups. For asymmetric, chiral molecules such as proteins, mirror and inversion operations are not allowed as they would not maintain the handedness of the structure. Thus there are 65 chiral space groups in which only rotational and screw axes are allowed symmetry operators (Hahn et al., 2005). For crystals containing chiral molecules the space groups can be grouped into seven different crystal systems: triclinic, monoclinic, orthorombic, tetragonal, trigonal, hexagonal and cubic. Taking into account primitive, face and body centered packing, 14 types of Bravais lattices are possible.

4.5.2 Solvent Content of Protein Crystals

Besides the polypeptide chain, protein crystals contain a considerable amount of disordered solvent, mostly water. Mathews first analyzed crystal structures for their solvent content, which led to the definition of the Matthews coefficient or Matthews volume (Matthews, 1968)

$$V_M = \frac{V_{UC}}{M \cdot z \cdot n} \tag{6}$$

where V_{UC} is the volume of the unit cell [Å^3], M the molecular mass [Da], z the number of molecules in the ASU and n the number of ASUs per unit cell. V_M generally lies between 1.5 and 6 Å^{3}Da^{-1}, with a mean around 2.5 Å^{3}Da^{-1}. Knowing the Matthews coefficient and assuming a density of 1.35 g/cm^3 for proteins, the fraction of the

crystal occupied by solvent can be computed:

$$\chi = 1 - \frac{1.23}{V_M} \tag{7}$$

A recent analysis has shown that a low solvent content (higher packing density) significantly correlates with increasing resolution of the crystal (Kantardjieff and Rupp, 2009). Taking this observation into account, a probability distribution for V_M at a given resolution limit (called Matthew probability) can be calculated that allows for a good estimation of the solvent content in proteins. The assessment of the number of molecules located in the unit cell is an important first step in molecular replacement (see section 4.6.7).

4.5.3 Crystallization of Proteins

Single protein crystals of sufficient size and quality are a basic prerequisite for X-ray structure analysis. Most crystallization techniques are based on gradually increasing the concentration of a precipitant which removes solvent and favors protein-protein interactions. A supersaturated protein solution reaches a metastable state where crystal growth but no nucleation is possible. By further increasing protein and precipitant concentration, the solution reaches an unstable phase in which nucleation can occur. Ideal conditions for crystallization are achieved when only a few crystallization nuclei form and the solution subsequently reaches the metastable phase again that allows for further growth of crystals. Salts, organic solvents and polymers are common precipitants used in the crystallization of proteins (Cudney et al., 1994; Carter Jr and Carter, 1979).

As it is usually not possible to predict the conditions in which a given protein crystallizes, a set of many different conditions is tested using so-called screens. Most often bifactorial grid screens and multifactorial sparse matrix screens are used for this purpose. In bifactorial screens, two parameters are varied, for example precipitant concentration and pH value, giving rise to a two-dimensional matrix. For multifactorial screens investigating the influence of many factors at once, it would be impractical to exhaustively test all variations. Therefore, modern multifactorial screens represent a sparse sampling of parameter space and are based on a statistical analysis of conditions that led to successful crystallization of different proteins in the past. Typically, three to four factors such as the type of buffer, pH value, precipitant type and concentration, and additional salts or small organic compounds, are varied (Jancarik and Kim, 1991).

A variety of different methods exist for protein crystallization, with vapor diffu-

sion and batch crystallization being the ones used most frequently (McPherson, 1999; Unge, 1999). For the work presented here, the vapor diffusion method was used exclusively, with hanging and sitting drop setups. Initial screening was performed in 96-well crystallization plates (Hampton Scientific and TTP Labtech) using the sitting drop method. For each drop, 0.1 - 0.4 μl of protein solution were mixed with an equal amount of reservoir buffer and equilibrated against 50 μl of reservoir. The plates were sealed with transparent viewseal (Greiner Bio-One). All screens were set up using Phoenix (Art Robbins Instruments) or Mosquito (TTP Labtech) liquid handling systems. The plates were incubated at 4 and 20 °C and periodically monitored by a RockImager System (Formulatrix). Optimization and refinement of initial crystallization conditions was performed in pregreased 24-well Linbro plates (Hampton Research) with the hanging drop method. Usually 1 μl of protein solution was mixed with the same volume of reservoir buffer on a small silanized glass plate and equilibrated against 500 μl of reservoir. The concentration gradient between the reservoir and the freshly set-up drops gets equilibrated through the vapor phase over time. Usually water vapor passes from the droplet to the reservoir, resulting in decreased volume and thereby increased concentrations of protein and precipitant in the drop, eventually leading to supersaturation and crystal formation.

Full-length antibodies have been shown to favor conditions of low ionic strength during crystallization (Harris et al., 2004). The same holds true for antibody fragments like Fabs. Based on the Low Ionic Strength Screen from Hampton Research, a 96-well screen named LISS was established for use with liquid handling systems.

4.5.4 Seeding Methods

One means to further optimize initially found crystallization conditions is the use of seeding techniques. Quite often only small crystals are initially obtained and growing larger ones can only be maintained by uncoupling nucleation and crystal growth by introducing preformed crystals to growth conditions (McPherson, 1999). In macroseeding, a crystal is transferred from the mother liquor in which it was grown to a new mother liquor that lacks protein. During this washing step, the crystal dissolutes on its surface as outer protein layers and contaminants that might be responsible for termination of growth are removed. The crystal is subsequently transferred to a new mother liquor containing protein that is less supersaturated than the initial liquor but sufficient to support crystal growth. For microseeding, a crystal is transferred to 100 μl of reservoir solution and smashed by vortexing with a seed bead (Hampton Research). The resulting seed stock is diluted and used to provide nuclei for new crystallization experiments. The third seeeding method is called streak seed-

ing. A stiff hair is inserted into a drop already containing crystals. By touching the crystal surface, small microcrystal fragments get attached to the hair which is then drawn across droplets of fresh mother liquor. These fragments may act as nucleation sites and grow into larger crystals.

For this work, microseeding and streak seeding was used to gain diffraction-quality crystals. Microseeding led to the identification of a number of crystallization conditions for the ternary complex of Fab2, Fab3H42 and PCSK9. Streak seeding allowed for the crystallization of aPCSK9-Fab1 in complex with a decameric PCSK9-peptide.

4.5.5 Cocrystallization and Soaking

There are two general approaches for crystallizing protein-ligand complexes. As protein crystals have a quite high content of disordered solvent of ∼25-80%, there are often discrete solvent channels present that pervade the crystal. If these channels are adjacent to the protein's binding site, it is possible for binding partners that are small enough to diffuse through the channels and bind to the binding site. This method of soaking can be used to produce heavy atom derivatives as well as complexes of proteins with small molecules or peptides. If the binding site is blocked by crystal contacts, or if soaking of a binding partner irreversibly alters and destroys the crystal, the only way to obtain complex crystals is cocrystallization.

Complexes of anti-Dabigatran Fabs with dabigatran or analogs were formed by incubation of the protein with 2.5 molar excess of the small molecule for 1 h at 4 °C. The Fab-hapten complexes were purified by SEC on a HiLoad-26/60-Superdex-75 prep grade (GE Healthcare Life Sciences) equilibrated and run in standard protein buffer. Prior to crystallization, the complexes were concentrated to 10 mg/ml and the buffer was changed to water.

For cocrystallization of aPCSK9-Fab1 with a decameric peptide the protein was incubated with a 2 molar excess of peptide for 1 h at 4 °C. The Fab-peptide complex was purified by SEC on a HiLoad-26/60-Superdex-75 prep grade (GE Healthcare Life Sciences) equilibrated and run in standard protein buffer. Prior to crystallization, the complex was concentrated to 10 mg/ml and the buffer was changed to crystallization buffer.

A ternary complex consisting of PCSK9, Fab2 and Fab3H42 was incubated with a 1.5 molar excess of each of the Fabs for 2 h at 4 °C. The ternary complex was purified and unbound Fabs were removed by SEC on a HiLoad-26/60-Superdex-200 prep grade (GE Healthcare Life Sciences) equilibrated and run in standard protein buffer. Prior to crystallization, the complex was concentrated to 10 mg/ml and the

buffer was changed to crystallization buffer.

4.5.6 Cryobuffer and Crystal Mounting

As protein crystals have a high solvent content, they must be protected from dehydration during X-ray data collection. A technique commonly used is to flash-cool the crystal in liquid nitrogen prior to collect data at 100 K. Collecting data at low temperatures is advantageous as it reduces thermal vibrations and minimizes radiation damage (Garman and Owen, 2005). Radiation damage can be observed as a decrease in diffraction intensity or resolution over time, as well as site-specific damage such a breaking of disulfide bonds or demethylation of methionines in the refined structure. It is frequently observed when using high intensity X-ray sources like third-generation synchrotrons. In addition, as thermal vibrations in crystals are reduced at low temperatures, the resolution to which the crystal diffracts will increase, an effect especially pronounced for high resolution datasets (Hope, 2006). Prior to flash-cooling, a cryoprotectant needs to be added, which prevents the water in the mother liquor to form ice crystals. These crystals would invariably destroy the diffraction quality of the crystal. Common cryoprotectants are glycerol, sucrose or low molecular weight PEGs.

A drop of mother liquor supplemented with 20% glycerol as cryoprotectant was placed on a glass plate. The crystals were scooped out of the mother liquor with a nylon loop and transferred to the cryobuffer for a period of $\sim$5 sec. The crystals were subsequently flash-cooled in liquid nitrogen.

4.6 X-ray Structure Solution

4.6.1 Theory of X-ray Diffraction

X-rays are high energy electromagnetic radiation with a wavelength of 1.0 to 0.01 nm (which corresponds to an energy of 1-125 keV). X-rays for protein crystallography can be generated in two ways. In laboratory X-ray sources, an anode of specific material (often copper) is bombarded by electrons. This results in electrons departing from core levels of the anode material. The remaining electron holes can be filled from the outer shells. During transition, the electrons emit the difference energy as X-ray photons with a wavelength characteristic of the anode's material. Alternatively, very intense X-rays can be generated with the help of a circular electron accelerator, a synchrotron (Rosenbaum and Holmes, 1971). Electrons (or other charged particles) that are deflected and thus experience an acceleration, emit radiation according to classical electrodynamics. In modern synchrotrons this deflection is accomplished by

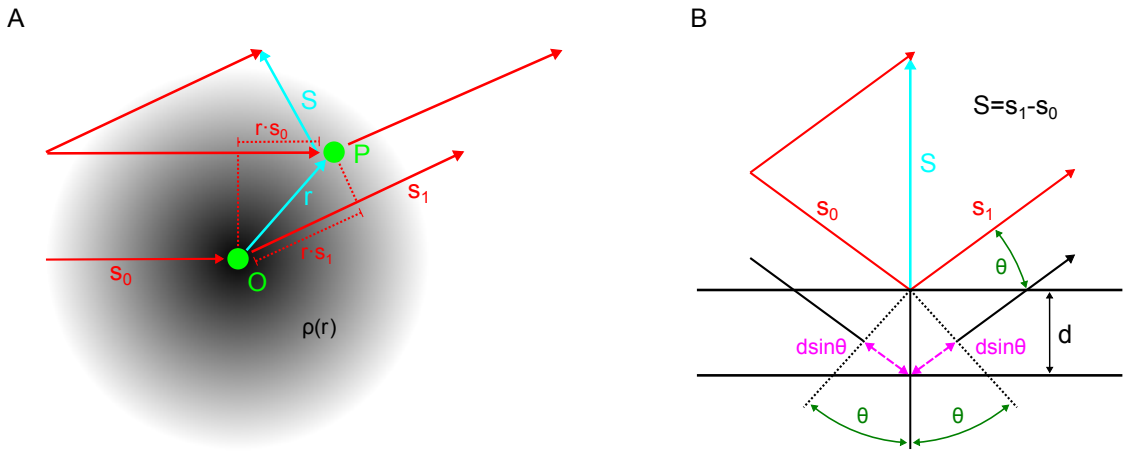


Figure 6: Diffraction of X-rays and Bragg's Interpretation. (A) Vector representation of a scattering event caused by two volume elements of an atom with electron density $\rho\mathbf{r}$). The relative path difference of the emanating waves is obtained from the scalar products: $\lambda\mathbf{r}\cdot\mathbf{s}_1 - \lambda\mathbf{r}\cdot\mathbf{s}_0$. **(B)** Graphical interpretation of the Bragg equation. Incoming X-rays are reflected on a set of lattice planes. The phase difference of the emanating waves is $2d\sin\theta$, where θ is the glancing angle. The phase difference $2d\sin\theta$ must equal a multiple of $n\lambda$ for constructive interference to result in a recordable reflection.

magnetic devices such as wigglers and undulators. Synchrotron radiation at first is polychromatic, but monochromators are used to select the desired wavelength. Besides the tunable wavelength, the high brilliance of the beam (defined as the number of photons per second that pass through an area of 1 mm^2 with a given divergence) is a big advantage of synchrotron radiation. X-rays (and electromagnetic radiation in general) interact with matter primarily through its oscillating electric field vector and only to a much lesser extent, the magnetic field vector. The electric field vector **E** (all vectors and complex quantities are presented in bold characters) is perpendicular to the propagation direction of the wave and oscillates with a frequency corresponding to the inverse of the wavelength of the X-rays. The dielectric polarizability α is the interaction constant relating the induced dipole moment **p** of matter to the electric field vector **E**:

$$\mathbf{p}=\alpha\mathbf{E} \tag{8}$$

In solid matter only the electrons of an atom are polarizable enough to interact with the electric field vector **E**, interaction with the atomic nuclei is neglible. When X-ray photons pass through matter, the electric field vector can induce oscillations in all the electrons within the photon's coherence length. The initial photon ceases to exist, but the oscillating electrons are able to emit an X-ray photon in a random direction, of the original energy and wavelength. This is called coherent scattering (Thomson scattering), as the scattering is coherent within the coherence length of each photon.

The wavelength does not change and a fixed phase relation exists between all partial waves emanating from all the electrons excited by an individual photon (Rupp, 2009). The effect of incoherent scattering (Compton scattering), where photons of lower energy are emitted, can be neglected. As the atomic distances in the crystal are of the same magnitude as the wavelength of the incident X-rays (∼1 Å), the scattered waves interfere with one another. In certain directions, the scattered waves interfere constructively, giving rise to detectable reflections, in all other direction they interfere destructively and cancel each other out. The symmetry of the resulting reflection pattern directly reflects the underlying symmetry of the crystal lattice, while its intensity distribution is determined by the arrangement of scattering matter within the repeating units of the crystal, in other words the electron density distribution. The diffraction process as well as the resulting path difference is shown in figure 6 on the preceding page. The incident wavelength is described by vector $\mathbf{s}_0$ with length $2\pi/\lambda$, the scattered wave by $\mathbf{s}_1$. The scattering vector $\mathbf{S}$ is the difference vector between incoming and scattered wave:

$$\mathbf{S} = \mathbf{s}_1 - \mathbf{s}_2 \tag{9}$$

The atom nucleus is placed at the origin O, the second diffracting volume element at point P with a distance $\mathbf{r}$. The path difference Δp of scattered waves emanating from the O and P is given as the difference of the scalar products:

$$\Delta p = (\lambda\mathbf{r} \cdot \mathbf{s}_1 - \lambda\mathbf{r} \cdot \mathbf{s}_0) = (\mathbf{s}_1 - \mathbf{s}_0) \cdot \mathbf{r}\lambda = \mathbf{S} \cdot \mathbf{r}\lambda \tag{10}$$

The path difference is converted to phase difference $\Delta\varphi$ in radians by multiplication with $2\pi/\lambda$

$$\Delta\varphi = 2\pi(\mathbf{s}_1 - \mathbf{s}_0) \cdot \mathbf{r} \tag{11}$$

integrating equation (9) into (11) leads to

$$\Delta\varphi = 2\pi \cdot \mathbf{r} \tag{12}$$

The resulting wave scattered from an atom can be described in exponential vector form with a relative phase $\exp(i\varphi)$ and a magnitude corresponding to its electron density $\rho(\mathbf{r})$. The atomic scattering factor $f_\mathbf{S}$ is obtained by integration over the entire

volume of the atom:

$$f_{\mathbf{S}} = \int_{\mathbf{r}}^{V(atom)} \rho(\mathbf{r}) \cdot \exp(2\pi i \mathbf{S}\mathbf{r}) d\mathbf{r} \tag{13}$$

A simplified description of scattering of X-rays in crystals has been introduced by William Bragg and allows diffraction to be described as the reflection of incident waves on a set of planes defined by the Miller indices *h, k* and *l* (see figure 7):

$$n\lambda = 2d_{hkl} \sin\theta_{hkl} \tag{14}$$

with *n* being integers, λ the wavelength of X-rays [Å], d_{hkl} the inter plane distance [Å] and θ_{hkl} is the "glancing angle" for reflection (*hkl*) [°]. Any plane of lattice points in a crystal can act as a Bragg plane. When equation (14) is fulfilled, the reflected waves will have a phase difference that is a integral multiple of their wavelength. The waves will constructively interfere with each other and a diffraction spot (therefore also called reflection) will be observed on the detector.

The scattering function of a molecule is obtained from the superposition of partial waves from all atoms of the molecule.

$$\mathbf{F}(\mathbf{S}) = \sum_{j=1}^{atoms} f_{\mathbf{S},j} \cdot \exp(2\pi i \mathbf{S}\mathbf{r}_j) \tag{15}$$

Scattering of a single molecule is practically impossible to detect. This is the reason, why crystals are used for X-ray structure determination. The periodic assembly of numerous molecules in the crystal lattice amplifies the scattered waves through constructive interference. The scattering function of the unit cell is obtained by integration of the electron density over the complete cell.

$$\mathbf{F}(\mathbf{S}) = \int_{\mathbf{r}}^{V(cell)} \rho(\mathbf{r}) \cdot \exp(2\pi i \mathbf{S}\mathbf{r}) d\mathbf{r} \tag{16}$$

Diffraction only occurs when the phase shift between incident and diffracted wave $(2\pi i \mathbf{S}\mathbf{r})$ is an integral multiple of 2π. When these so called Laue conditions are fulfilled, constructive interference can occur.

$$\mathbf{a} \cdot \mathbf{S} = h$$

$$\begin{aligned} \mathbf{b} \cdot \mathbf{S} &= k \\ \mathbf{c} \cdot \mathbf{S} &= l \end{aligned} \tag{17}$$

with *h, k, l* being integers. These integers are called Miller indices and describe a set of lattice planes in real space that can be also described with the lattice vector **h**. The miller indices are also used to identify the reflection that results from a given set of lattice planes. If the positional vector **r** is described by the fractional cell coordinates *x, y* and *z*, then taking into account equation (17), the structure factor can be described as

$$\begin{aligned} \mathbf{F}(\mathbf{h}) &= V \int_{x=0}^{1} \int_{y=0}^{1} \int_{z=0}^{1} \rho(xyz) \cdot \exp(2\pi i(hx + ky + lz) dx\, dy\, dz \\ &= V \int_{\mathbf{x}=0}^{1} \rho(\mathbf{x}) \cdot \exp(2\pi i \mathbf{h}\mathbf{x}) d\mathbf{x} \end{aligned} \tag{18}$$

which is in the form of a Fourier integral. According to the Fourier integral theorem, one can perform a back-transformation of equation (18) from reciprocal space over all reflections **h** to real space:

$$\rho(\mathbf{x}) = \frac{1}{V} \int_{\mathbf{h}=-\infty}^{+\infty} \mathbf{F}(\mathbf{h}) \cdot \exp(-2\pi i \mathbf{h}\mathbf{x}) d\mathbf{h} \tag{19}$$

The scattering in reciprocal space, the complex structure factor **F** corresponds to the electron density of the unit cell in real space the sum of all atomic scattering factors *(f_j)* within the unit cell. From equations (18) and (19) it becomes obvious, that electron density $\rho(\mathbf{x})$ in real space can be computed from the structure factors $\mathbf{F}(\mathbf{h})$ in reciprocal space and vice versa. According to the Laue conditions (see equation (17)) diffraction can only occur at discrete positions, which transforms the Fourier integration into a summation:

$$\rho(x, y, z) = \frac{1}{V} \sum_{h=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} \mathbf{F}(hkl) \cdot \exp(-2\pi i(hx + ky + lz)) \tag{20}$$

and

$$\mathbf{F}(hkl) = V \sum_{x=0}^{1} \sum_{y=0}^{1} \sum_{z=0}^{1} \rho(x,y,z) \cdot \exp(2\pi i(hx + ky + lz)) \tag{21}$$

The structure factor $\mathbf{F}(\mathbf{h})$ is a complex number defined by its amplitude $F(\mathbf{h})$ and phase $\alpha(\mathbf{h})$ (see equation (22)). The amplitudes can be computed from the intensities of the diffraction spots $I(\mathbf{h})$ (see equation (23)), while the phases cannot be determined directly. Nonetheless the phases can be obtained with experimental methods or computed by molecular replacement (see section 4.6.7)

$$\mathbf{F}(\mathbf{h}) = F(\mathbf{h}) \cdot \exp(i\alpha(\mathbf{h})) \tag{22}$$

$$F(\mathbf{h}) = \sqrt{I(\mathbf{h})} \tag{23}$$

4.6.2 Reciprocal Lattice and Ewald Sphere

As explained above, diffraction on a set of lattice planes (*hkl*) gives rise to a reflection *hkl*. The real space lattice planes *(hkl)* are represented by a vector $\mathbf{d}^*_{hkl}$ in reciprocal space that extends from the origin of the reciprocal lattice to a reciprocal lattice point *hkl*. The reciprocal lattice points can be described as integral multiples of the basis vectors, which are defined as $\mathbf{a}* = \mathbf{S}_{100}$, $\mathbf{b}* = \mathbf{S}_{010}$ and $\mathbf{c}* = \mathbf{S}_{001}$. Taking into account the Laue conditions, equation 24, that relates real and reciprocal space, can be obtained

$$S = h\frac{\mathbf{b}\times\mathbf{c}}{\mathbf{a}(\mathbf{b}\times\mathbf{c})} + k\frac{\mathbf{c}\times\mathbf{a}}{\mathbf{a}(\mathbf{b}\times\mathbf{c})} + l\frac{\mathbf{a}\times\mathbf{b}}{\mathbf{a}(\mathbf{b}\times\mathbf{c})} = h\mathbf{a}* + k\mathbf{b}* + l\mathbf{c}* \tag{24}$$

Bragg's law (equation (14)) can be rewritten so that its relation to the reciprocal lattice becomes obvious:

$$\mathbf{d}^*_{hkl} = 1/d_{hkl} = 2\sin\theta/n\lambda \tag{25}$$

Paul Ewald introduced a geometrical construction that illustrates the relation between real and reciprocal lattice and visualizes under which circumstances Bragg reflections occur (Ewald, 1921). The crystal is situated in the center of the Ewald sphere (with radius $1/\lambda = \mathbf{s}_1/2\pi$). The origin of the reciprocal lattice is defined as the intersection of the incident beam vector with the Ewald sphere. Bragg reflection occurs when a reciprocal lattice point, i.e. the end point of a diffraction vector $\mathbf{d}^*_{hkl}$,

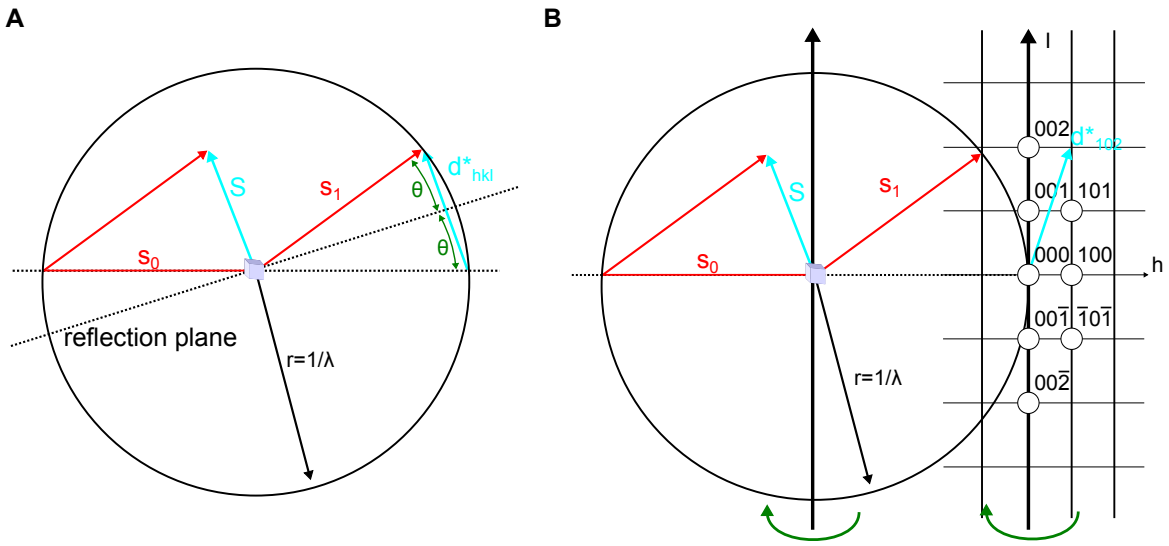


Figure 7: Ewald Sphere and Incidence of X-ray Reflections. **(A)** The incident wave $\mathbf{s}_0$ is diffracted at a Bragg reflection plane, the scattered wave has a relative angle of 2ϑ. Ewald's sphere is drawn with a radius of $1/\lambda$ around the origin of the crystal lattice. **(B)** The intersection of the incident wave with the Ewald sphere is the origin of the reciprocal lattice. Rotation of the crystal is translated into rotation of the reciprocal lattice. Reflections occur, when a reciprocal lattice point crosses the Ewald sphere, as only then the Laue equations are fulfilled.

intersects with the Ewald sphere. Under these conditions the vector is perpendicular to the reflecting lattice planes and coincides with scattering vector **S**. This means the condition expressed in equation (25) and thus Bragg's law is met. As immediately visible from figure 7, at a fixed orientation only few spots of the reciprocal lattice will lie on the Ewald sphere and thus give rise to few observable reflections. By rotating the crystal, the reciprocal lattice is rotated the same way, which brings additional reciprocal lattice points to intersect with the Ewald sphere.

In the absence of anomalous scattering contributions, the reciprocal lattice contains a center of inversion which means that the intensity I of (the centrosymmetrically opposed conjugate) reflections hkl and $\overline{hkl}$ have the same intensity. There is no difference in the intensity of X-rays scattered from the top or bottom side of a set of lattice planes. This observation is described in equation (26) and called Friedel's law (the reflection pairs are called Friedel mates).

$$\begin{aligned} |\mathbf{F}(hkl)|^2 &= \left|F^*(\overline{hkl})\right|^2 \\ I(hkl) &= I(\overline{hkl}) \end{aligned} \tag{26}$$

4.6.3 Temperature Factors

The structure factor **F(h)** can be described as the sum of all n atomic form factors f_j within the unit cell. f_j decreases with scattering angle and is described by Cromer-

Mann approximation. In reality, atoms in a crystal lattice are not fixed, but vibrate around their mean position. Thermal vibration is more pronounced at higher temperatures. In addition, atoms can be displaced in the crystal because of disorder. Both effects lead to an additional decrease in intensity of the resulting reflections, especially at high glancing angles, i.e. high resolutions (Debye, 1914). This is accounted for by a correctional term $\exp(-B_{iso}|\mathbf{S}|^2/4) = \exp(-B_{iso}(\sin\theta/\lambda)^2)$, the Debye-Waller factor which is also called temperature factor *B* (see equation (27)). It is proportional to the isotropic mean square displacement $\langle u_{iso}^2 \rangle$ of an atom from its resting position as described in equation (28). Isotropic displacement is only an approximation to the actually anisotropic displacement in real crystals. As the determination of anisotropic displacement significantly increases the number of parameters, *B*-factors can be parametrized this way only for high resolution structures or with the help of TLS-parameterization (see section 4.6.9).

$$F(\mathbf{h}) = \sum_{j=1}^{n} f_j^0 \exp(2\pi i \mathbf{S}\mathbf{r}_j) \exp\left(-B_{iso}|\mathbf{S}|^2/4\right) \tag{27}$$

where f_j^0 is the form factor of atom *i* in resting position and B_{iso} the temperature factor. B_{iso} is defined as

$$B_{iso} = 8\pi^2 \left\langle u_{iso}^2 \right\rangle \tag{28}$$

Determination of Scale and Temperature Factors The recorded experimental data need to be converted to a common scale with the theoretical model data by two factors. The scale factor *k*, which brings structure factor amplitudes onto an absolute scale, as well as the overall temperature factor B_{iso} that is specific for the respective protein structure, can both be determined with the help of a Wilson plot (Wilson, 1949). All measured intensity data are divided into resolution bins each containing an equal number of reflections. The absolute intensity $\langle I_{abs} \rangle$ can be calculated from the sum of all scattering factors at rest:

$$\langle I_{abs} \rangle = \sum_{j=1}^{atoms} \left(f_j^0\right)^2 \tag{29}$$

A graph is drawn of the natural logarithm of the quotient of the mean observed intensity $\langle I_{obs} \rangle$ and $\langle I_{abs} \rangle$ in each resolution bin against $(\sin\theta/\lambda)^2$. As a consequence of secondary structure motifs, the atomic distances in proteins are not equally populated, especially at low resolutions. These Debye effects cause deviations from linearity in

the Wilson plot at resolutions above 3.2 Å. The scale factor *k* and the temperature factor B_{iso} can be readily obtained from linear regression in the range between highest resolution and 3.2 Å (see equation (30)), when the approximate composition of the unit cell is known. The intercept of the regression line determines $\ln k$, the slope of the regression line determines the overall *B*-factor. Both parameters, *k* and B_{iso}, are later adjusted during structural refinement.

$$\ln \frac{\langle I_{obs} \rangle}{\langle I_{abs} \rangle} = \ln k - 2B_{iso} \left(\frac{\sin \theta}{\lambda} \right)^2 \tag{30}$$

4.6.4 Patterson Function

The Patterson function is defined at any point **u** by a convolution integral of the electron density over the whole unit cell in real space *R* (Patterson, 1934). For the Patterson function the relative coordinates **u** (*u, v, w*) are used to prevent confusion with coordinates **r** (defined by *x, y* and *z*). The dimensions of the Patterson cell are equal to those of the unit cell.

$$P(\mathbf{u}) = \int_R \rho(\mathbf{r}) \rho(\mathbf{r} + \mathbf{u}) d\mathbf{r} \tag{31}$$

can also be written as

$$P(\mathbf{u}) = \frac{1}{V} \sum_{\mathbf{h}} F_{\mathbf{h}}^2 \cos 2\pi(\mathbf{hu}) \tag{32}$$

The Patterson function can be computed directly from the measured amplitudes of the structure factors without knowledge of their phases. The product of electron densities at the locations **r** and $\mathbf{r}+\mathbf{u}$ is only non-zero if **u** is the difference vector between two atoms. The Function can thus be interpreted as a summation of all vectors between all atoms in the unit cell. With N atoms in the unit cell, there will be N^2 peaks in the Patterson function, of which $N(N+1)$ will be non-origin peaks that describe inter-atomic distances. The height of each distance vector peak is proportional to the product of the electron densities of both atoms. As there are two distance vectors ($\overrightarrow{AB}$ and $\overrightarrow{BA}$) between the points A and B, the Patterson function is centrosymmetric in addition to the symmetry of the crystal.

From a map constructed from the Patterson function, the positions of atoms can be determined without phase information for structures consisting of only few atoms within the unit cell. For protein structures with several thousand atoms, the Patterson map gets too crowded and is no longer directly interpretable. Nonetheless, these

maps are used for the determination of marker atom positions in experimental phasing from isomorphous difference data, as well as anomalously scattering atoms in dispersive difference data. As explained in section 4.6.7, the Patterson function is invaluable for the search model rotation during molecular replacement.

4.6.5 Data Collection

The accurate assessment of the positions and the intensity of the Bragg reflections is the basis for X-ray structure solution. Measurement of the reflections was performed according to the rotation method (Arndt et al., 1973), which involves the rotation of the crystal around a single axis in small increments. A diffraction image is recorded for each increment of rotation (for this thesis, increments of 0.5 ° were chosen). The resulting frame contains positional coordinates of individual picture elements, intensities and information on the detector geometry. The reflections were recorded using X-ray area detectors such as mar345 Image Plate (Mar Research) and Saturn 944 HG CCD (Rigaku Corporation) inhouse. At the synchrotron beamline a PILATUS 6M hybrid pixel array detector (DECTRIS Ltd.) was used because of its superior signal-to-noise ratio and the low read-out time of 5 ms that allows for data collection in a continuous mode (Henrich et al., 2009). The X-rays were generated with a MicroMax-007 HF rotating copper anode generator (Rigaku) in house, that was operated at a current of 30 mA and a voltage of 40 kV. The characteristic Cu K_α radiation with a wavelength of 1.54 Å was selected with the help of a multilayer graphite monochromator and the divergence was reduced with a collimator to 0.5 mm.

Prior to measurements on a synchrotron beamline, all crystals were tested at one of two home diffractometers. As experimental setup a typical goniometer system adjustable at four angles (4-circle goniostat) was used. The desired orientation of the crystal was adjusted with the angles ω, φ and χ and for data collection the crystal was rotated around the φ-axis in 0.5° increments. Frames were collected with an exposure time of 5 min in case of the image plate and 30 sec for the CCD detector. The detector distance was adjusted to record a maximal resolution while allowing for the recording of individual separated spots. During data collection, all crystals were maintained at cryogenic temperatures of 100 K with a Cobra cooler (Oxford Cryosystems). The first 10 frames recorded were used to generate a strategy for data collection with XPLAN (Kabsch, 2010) as implemented in an in-house shell script.

Synchrotron data were collected at the X06SA Swiss Light Source (SLS) beamline at Paul Scherrer Institute (Villigen, Switzerland). X-ray diffraction data were recorded at cryogenic temperature (100K) using 0.91 Å synchrotron radiation. High resolution datasets consisting of 360 images (0.5°/frame, 0.5 seconds exposure) from single

crystals were collected on a PILATUS 6M detector.

4.6.6 Processing, Scaling and Reduction of Data

All recorded data were processed using the autoPROC toolbox (Vonrhein et al., 2011), which provides a framework for several individual modules such as SCALA (Evans, 1997), XDS (Kabsch, 2010) and POINTLESS (Evans, 2005).

The first step in data processing is to determine the dimensions and orientation of a primitive crystal lattice relative to the laboratory coordinate system. The centroids of the strongest reflections are used to index the crystal lattice which allows for subsequent determination of the Laue symmetry. As many reflections are only partially recorded on a single frame, so-called batches have to be calculated, i.e. three-dimensional spot profiles from several consecutive frames. These batches usually comprise about 5 to 10° of rotation. In the actual integration step, profile fitting is performed. All reflections are integrated by fitting them to the best available three-dimensional profile. In the following postrefinement, reflections that lie close to their predicted positions are used to refine geometric parameters, such as cell parameters, detector distance and orientation. The last two steps are iteratively repeated until convergence of the data. The resulting file contains the so-called raw reflection intensities (Leslie, 2005). These data, together with their standard deviations, are then corrected for absorption and Lorentz and polarization effects for the given beam geometry and crystal orientation. In addition to the spatial integration, it must also be considered that the reflections were recorded over a period of time which demands for temporal scaling. This allows for the correction of effects such as fluctuating beam intensity, anisotropic diffraction of the crystal and loss of diffraction power caused by radiation damage, which is often seen at synchrotron beam lines. Intensities are thus scaled within the data set in order to reduce the mean variation between multiple measurements of equivalent reflections (Rupp, 2009). At this stage, it is already possible to determine the lowest symmetry space group by comparison of the intensities with that of a given symmetry-related space group and by analyzing the systemic absence caused by lattice centering and occurrence of screw axes, respectively. As identical or symmetry related reflections are recorded multiple times during data collection, these reflections are merged into a unique data set representing the asymmetric unit of the reciprocal space for a given space group.

There are several indicators for the quality assessment of the recorded data. As a consequence of atomic scattering factor fall-off and of atomic and lattice displacement, the intensity of reflections decreases at higher resolutions. Data statistics are therefore often described as binned in resolution shells that contain an equal num-

ber of reflections. The decrease in intensity manifests in a lower signal-to-noise ratio $\langle |I| / \sigma(I) \rangle$ for high-resolution data, that is expressed in equation (33) with *N* being all reflections in a given resolution shell and $\sigma(I)$ the standard deviation of the intensity.

$$\langle |I| / \sigma(I) \rangle = \frac{1}{N} \sum_{\mathbf{h}}^{N} \frac{\left| I_{(\mathbf{h})} \right|}{\sigma(I_{(\mathbf{h})})} \tag{33}$$

Another quality indicator for collected data is the linear merging *R*-value that compares the intensity of identical or symmetry related reflections

$$R_{merge} = \frac{\sum_{\mathbf{h}} \sum_{i=1}^{N} \left| I_{(\mathbf{h})i} - \bar{I}_{(\mathbf{h})} \right|}{\sum_{\mathbf{h}} \sum_{i=1}^{N} I_{(\mathbf{h})i}} \tag{34}$$

with *N* being all redundant observations for reflection **h** and $\bar{I}_{(\mathbf{h})}$ the averaged intensity of each reflection. As R_{merge} does not account for the redundancy of the data, higher redundancy results in an increased *R*-factor. To overcome this limitation, a more robust criterion was introduced where each reflection is weighted according to its multiplicity (Diederichs et al., 1997). This value is called redundancy-independent merging *R*-value (R_{rim} or R_{meas})

$$R_{rim} = \frac{\sum_{\mathbf{h}} \sqrt{\left(\frac{N}{N-1}\right)} \sum_{i=1}^{N} \left| I_{(\mathbf{h})i} - \bar{I}_{(\mathbf{h})} \right|}{\sum_{\mathbf{h}} \sum_{i=1}^{N} I_{(\mathbf{h})i}} \tag{35}$$

with *N* being the number of independent measurements of reflection **h**.

4.6.7 Phasing by Molecular Replacement

A major obstacle in solving crystal structures of macromolecules is the inability to directly determine the phases of the diffracted X-rays, the crystallographic phase problem (see section 4.6.1). Nevertheless, there are ways to gain information about the phase of each reflection. Direct methods use statistical relationships between certain sets of structure factors. As these relationships become weaker with increasing number of atoms, direct methods are limited to small proteins and atomic resolution data (better than 1.2 Å) and thus can not be used to solve large macromolecules. Experimental phasing is based on the determination of the substructure of marker atoms, that provide a source of electronic differences relative to isomorphous refer-

ence structures (Rupp, 2009). This can be accomplished by isomorphous replacement where heavy atom derivatives of the crystals are prepared by soaking (Green et al., 1954). Alternatively anomalous dispersion, which describes the effects of electronic resonance at certain energies can be used if dispersive atoms in the crystal and a tunable X-ray source, such as a synchrotron, are available.

If there are structures of homologous proteins available, molecular replacement (MR), a term first introduced by Rossmann, can be used to obtain information about the phases of a unknown crystal structure (Rossmann, 1972). MR can also be used, if proteins crystallize in different space groups or with a ligand bound, with both cases often leading to slight conformational rearrangements. Statistical analysis has shown that proteins that share a sequence identity of at least 30% often share the same fold and display a relatively small root mean square deviation (rmsd) of less than 1.5 Å (Sander and Schneider, 1991) for their Cα-atoms, which often suffices to identify a starting search model for MR. With recent advances in including protein structure modeling algorithms, it was possible to solve structures that share only 20% sequence identity with the best starting models (DiMaio et al., 2011). Fortunately, the number of possible protein folds is restricted to 500 to 1000 (Brenner et al., 1997) and with the increasing number of protein structures solved, it becomes more and more likely to find a homologue structure close enough to the protein to be solved. Molecular replacement was the solely phasing technique used during this work. The initial search models were selected from a local DSSP-server according to their sequence identity. For Fabs, the individual domains were used as separate search models. At positions where the search model sequence deviated from the actual protein in the crystal, the model was truncated to its Cβ-atoms with the program CHAINSAW (Stein, 2008). The highly divers CDR loops of Fabs, especially CDR:H3, were completely removed from the search models if they differed from the protein in sequence or length.

The principle of molecular replacement is based on the placement of a search model in the unit cell in such a way that the calculated structure factor amplitudes match the measured reflection amplitudes. When the search model is correctly oriented and positioned, the phases for each structure factor can be calculated according to equation (27). With the measured structure factor amplitudes it is then possible to calculate an electron density map and start model building and refinement. The placement of the search model is a six-dimensional problem which can be simplified by splitting it into a rotation function (RF) and a translation function (TF) search (Rossmann and Blow, 1962). The translation and the rotation search can be implemented as Patterson correlation functions or as maximum-likelihood (ML) functions (Bricogne, 1992). For both methods, the basis for a successful search is a high degree of structural homology between search and target molecules as well as a high

quality of the diffraction data. Furthermore, the number of monomers in the cell and the complexity of the space group play a role. For this work the maximum-likelihood program PHASER (McCoy et al., 2007) was used for molecular replacement. The likelihood of the models is generated on a grid of angles or positions and the model with the highest likelihood is selected (McCoy, 2004). Maximum-likelihood RF and TF can include partial structure information. This is especially useful when solving protein complexes (e.g. antibody-antigen complexes) or proteins with flexible domains (e.g. antibody fragments that have a flexible elbow angle). In such cases a "tree search with pruning" strategy is used, where partial solutions increase the signal-to-noise ratio of the search for the following components (McCoy, 2006). A quality indicator that allows to judge if a given MR solution is correct is the Z-score. It compares the log-likelihood gain (LLG) from RF and TF search with the LLG of a set of random rotations and translations. The mean and the rmsd from the mean are computed from the random set, then the Z-score for a search peak is defined as its LLG minus the mean, all divided by the rmsd, i.e. the number of standard deviations above (or below) the mean. A Z-score of 6 or higher can be expected for a correct solution.

4.6.8 Model Building

With phases calculated from MR and the measured amplitudes of the reflections, the electron density map of the crystal can be constructed. Given good quality data, the map can be interpreted and a protein model can be built into the density. In this work the program COOT (Emsley and Cowtan, 2004) was used for model building. The MR solution was checked for atom clashes and crystal packing. As during MR phasing solely rigid body rotations and translations are performed, the resulting atom coordinates are first approximations to their real positions in the crystal. One round of refinement (see section 4.6.9) was thus performed prior to model building, which resulted in an electron density map that was much easier to interpret. The truncated or missing amino acids and ligands were built where the electron density was unambiguous. After each round of rebuilding, the model was refined in reciprocal space and the resulting electron density inspected, whereby the following types of electron density maps were used:

($F_{obs} - F_{calc}$) electron density map Difference Fourier maps with the coefficients $(F_{obs} - F_{calc}) \exp(i\alpha_{calc})$ show differences between the model and the true structure. Missing parts of the model appear as positive and wrongly positioned as negative difference density. The differences occur with half of the actual electron density and usually become significant at a contour level of three standard deviations (3 σ) of

the mean electron density. ($F_{obs} - F_{calc}$) density maps are the basis for correction and rebuilding of the protein model as well as for the placement of water molecules or small molecule ligands.

($2F_{obs} - F_{calc}$) electron density map An electron density map calculated with the coefficient $(2F_{obs} - F_{calc})\exp(i\alpha_{calc})$ can be described as a superposition of an $(F_{obs})\exp(i\alpha_{calc})$ and an $(F_{obs} - F_{calc})\exp(i\alpha_{calc})$ electron density map. These maps are usually viewed at a contour level between 1.0 and 1.5 σ and show the modeled protein. ($F_{obs} - F_{calc}$) and ($2F_{obs} - F_{calc}$) electron density maps sometimes are strongly influenced by model bias. Therefore solely σ_A-weighted electron density maps were used for this work.

σ_A-weighted ($mF_{obs} - DF_{calc}$) and ($2mF_{obs} - DF_{calc}$) electron density maps There will be errors in both the phase and the amplitude of the calculated structure factors, as well as in the amplitude of the observed structure factors. Errors for F_{calc} arise if parts of the model are not yet built or incorrectly placed. Errors in the measured reflection intensities propagate into errors for F_{obs}. To address this problem, the calculation of structure factor amplitudes must be weighted accordingly. Therefore, corrections are introduced that account for errors in the model like incompleteness and disorder (*D*), as well as for errors in scaling and intensity measurements (*m*) (Read, 1986, 1990). Maps constructed from the maximum-likelihood Fourier coefficients $(2mF_{obs} - DF_{calc})\exp(i\alpha_{calc})$ for acentric, and $mF_{obs} \cdot \exp(i\alpha_{calc})$ for centric reflections, incorporate these errors in the so-called σ_A variance term (from which *m* and *D* can be calculated). The resulting maps provide minimally biased electron density (Read, 2006) and were used for model building.

4.6.9 Refinement

After model building in real space, the model still contains a variety of deviations concerning atom positions and displacement factors. These errors and therefore the deviation of the model from the measured diffraction data can be minimized with reciprocal space refinement. The parameters describing the model, such as atom coordinates, *B*-factors and overall parameters like scale factor and overall *B*-factor are refined against the experimental data, to obtain a best fit between the computed model structure factor amplitudes and the observed structure factor amplitudes. This fit between model parameters and diffraction data is monitored as a linear residual

(*R*-factor):

$$R = \frac{\sum_{\mathbf{h}} |F_{obs} - F_{calc}|}{\sum_{\mathbf{h}} F_{obs}} \tag{36}$$

In general, models can only be refined when the number of observations exceeds the number of parameters. Initial observations are the recorded reflections which are limited in resolution and completeness. The number of observations can be increased by introducing universally valid observations that reflect chemical knowledge about protein structures. These can be restraints for the lengths and angles of atomic bonds and dihedral angles (Engh and Huber, 1991) and for the *B*-factors of adjacent atoms. The presence of non-crystallographic symmetry (NCS) can also be exploited, as multiple copies of a protein in the asymmetric unit will usually have very similar conformations, especially in their backbone geometry. In addition, there are constraints like the sum of partial occupancies of atoms being limited to 1 or the sum of all bond angles in a planar group with three ligands being 360°. A reasonable model will have few stereochemical violations, low fluctuations in *B*-factors and will explain crystallographic, biophysical and biochemical data with a minimal number of parameters. During refinement, the danger of over-fitting always persists. To overcome this problem the method of statistical cross-validation and the corresponding free *R*-value (R_{free}), with *k* being a resolution-dependent scale factor (see equation (37)), was introduced (Bruenger, 1992). Prior to the first refinement, the data are divided into a large "working" set (comprising ~95% of the reflections) and a "test" set (comprising the remaining 5%). The data in the working set are used for refinement, the respective *R*-value is then referred to as R_{work}. The agreement between the model and data is then subsequently computed separately for the working (R_{work}) and test set (R_{free}). The free *R*-value measures the degree to which the model predicts the diffraction data for the test set which was excluded from modeling and refinement. It is related to the mean phase error and thus a statistical means to assess the improvement of a model during modeling and refinement (Kleywegt et al., 1996). In case of over-fitting of the model, R_{free} will increase although R_{work} might be lowered.

$$R_{free} = \frac{\sum_{\mathbf{h} \in free} |F_{obs} - kF_{calc}|}{\sum_{\mathbf{h} \in free} |F_{obs}|} \tag{37}$$

The refinement program autoBUSTER that is based on the principle of maximum likelihood was used for this work. ML allows to account for errors in both the model

and the recorded data. The probability of a model being correct is calculated by Bayesian inference. The best model is the one that has a large likelihood to describe the measured data (called P(Xray)) while having a high prior probability in respect of chemical restraints (P(chemistry)) (Bricogne and Irwin, 1996).

$$P(\text{refinement}) = P(\text{chemistry}) \times P(\text{Xray}) \tag{38}$$

During refinement P(chemistry) and P(Xray) are weighted depending on the resolution of the dataset. The higher the resolution and quality of the recorded reflections, the higher P(Xray) is weighted. For low resolution datasets, P(chemistry) is given greater weight which corresponds to more strict geometry restraints. In practice, it is always necessary to overweight P(Xray) for the refinement to converge (McCoy, 2004). The optimal choice of weights is that which gives the global maximum of the R_{free} likelihood, provided the refinement had converged at the global maximum of the R_{work} likelihood (Tickle et al., 1998).

Model Parameterization: Coordinates In general, a protein model can be described by the coordinates of the individual atoms (*x*, *y* and *z*) and the displacement factor *B*. This results in 4 parameters for each atom, in case of anisotropic B-factors 10 parameters. In addition, there are overall model parameters such as anisotropic scale factors and overall *B*-factors. The lower the resolution of the dataset, the lower the number of observables. To allow refinement, the number of model parameters must therefore also be lowered, which can be accomplished in several ways. In case of a very low resolution data, it is worthwhile to parametrize the individual proteins as rigid bodies. The internal structure of the protein is regarded to be rigid, so there are only 6 parameters for each protein, three to describe its orientation and three to describe its position. This method has a high radius of convergence and can thus be performed at low resolutions. In fact it is also one of the final steps during MR phasing with PHASER (see section 4.6.7). A less drastical reduction of model parameters is used for torsion angle refinement. As the lengths and angles of atomic bonds display only slight variations (Engh and Huber, 1991), they can be constrained to their mean values. This allows to describe the protein's backbone geometry by two main chain torsion angles for each amino acid. Regarding an average residue which has about eight atoms and five torsion angles, only five parameters need to be refined with this method. Compared to the individual positional refinement, where 24 parameters have to be described, this corresponds to an almost five-fold decrease in the number of parameters. In case of non-crystallographic symmetry (NCS), i.e. when multiple copies of the same protein (complex) are present in the crystal and

their relations can not be described by proper space group symmetry operations, the individual proteins usually share very similar backbone conformations for the bulk of the protein, whereas for the sidechain orientations there is often a greater variability. Their backbone geometries can be described by a set of parameters that is valid for all NCS related molecules, which significantly decreases the overall number of parameters and allows for NCS-restrained refinement. Local differences between the NCS related molecules are accounted for by pruning them from the NCS-restraints in autoBUSTER with the “autoncs”-flag.

Model Parameterization: *B*-factors As mentioned in section 4.6.3, the atomic displacement parameter (*B*-factor) describes the displacement of an atom from its mean position. It includes thermal vibration about the equilibrium position, as well as static and dynamic disorder. The simplest description is achieved with an overall isotropic *B*-factor for the whole structure, which is a very rough approximation of real atomic displacement. Individual isotropic *B*-factor parameterization can be used for resolutions of 2.5 Å or better. The most detailed description of atomic displacement is achieved with individual anisotropic *B*-factors. Here the isotropic displacement vector is replaced by an anisotropic displacement tensor that is described by 6 parameters for each atom. Because of the resulting increase in the number of parameters, anisotropic *B*-factor refinement is only possible for high resolution structures of 1.4 Å or better.

A well established approximation of anisotropic displacement for medium resolution datasets are translation-libration-screw (TLS) parameters (Schomaker and Trueblood, 1968; Winn et al., 2001). The symmetric tensors **T** and **L** describe the translation in units of Å^2 and the rotational components in units of rad^2. The third, asymmetric tensor **S** describes the correlation between the translation and the rotation movement. Each TLS group is treated as a rigid body and its anisotropic thermal displacement is described with 20 parameters. This allows for the anisotropic parameterization of atomic displacement at medium resolutions. A TLS group can consist of individual amino acid sidechains, flexible loops, domains, amino acid chains or whole proteins. In the autoBUSTER package, one TLS group is defined for each protein chain in the crystal, as this was shown for most crystal structures to result in an optimal decrease of R_{free} while using a minimum number of TLS groups. Only for the ternary complex of PCSK9 the four immunoglobulin domains of Fab2 were treated as individual TLS groups for reasons that will be explained later. To keep the number of initial parameters low, TLS refinement was performed at a late stage of model building, which lead to an additional decrease in R_{free} in all structures.

Solvent and Ligand Placement Water molecules were placed automatically with the "WaterUpdatePkmaps" macro as implemented in autoBUSTER. Criteria for placement were the presence of spherical difference density with a suitable donor/acceptor nearby. New waters were only placed close to existing hydrogen partners, wrongly placed waters were deleted automatically. Water molecules were also built and deleted manually in COOT.

The small molecule dabigatran and analogous compounds were manually fit into ($mF_{obs} - DF_{calc}$) difference density maps using COOT. 2D structures for all small molecule ligands were obtained from Boehringer Ingelheim's compound data base (CDB) and ligand dictionaries were created with GRADE as implemented in an in-house shell script. GRADE uses the programs libcheck (Vagin et al., 1998) to produce an initial dictionary, mogul (Bruno et al., 2004) to obtain ideal bond lengths, angles and torsions from similar structures in the CSD (Cambridge Structural Database) and finally performs semi empirical quantum mechanics calculations for geometry optimization as implemented in the fdynamo library (Field et al., 2000).

4.6.10 Structure Validation

The protein model was permanently validated during building and refinement. There are global and local structure quality indicators, which describe the overall model accuracy and loops, single amino acids or atoms, respectively. For proper structure validation, only those parameters should be analyzed that were not part of the refinement target function. Validation was performed during model building as implemented in COOT and after each cycle of reciprocal space refinement with a subsequent MOLPROBITY analysis (Chen et al., 2009). The MOLPROBITY output was used for manual correction of regions with low geometric quality in COOT.

Global indicators for the agreement of the modeled structure with the measured reflection intensities are the residual values R_{work} and R_{free} (see section 4.6.9). The combination of both residual factors is a valuable tool to describe the overall deviation of model and data, but they can not be used for local structure validation.

Local quality indicators focus on the stereochemistry of individual residues or atoms. The dihedral angles φ and ψ of the protein backbone can adopt a variety of conformations depending on the type of amino acid. In a Ramachandran plot these dihedral angles are visualized for every amino acid of the protein (Ramachandran et al., 1963). Amino acid outliers in the Ramachandran plot indicate energetically unfavored or chemically impossible backbone conformations. Favored and allowed regions for the Ramachandran plot are defined in MOLPROBITY, which enclose 98% and 99.95% of the conformations seen in a reference set of protein structures, re-

spectively. An updated version of the "penultimate rotamer library" (Lovell et al., 2000) was used for the analysis of amino acid side chain torsion angles. Rotamers that occur with a frequency of less than 1% in this library are designated as "poor rotamers" and were changed to more frequently occurring rotamers whenever electron density and local geometry allowed for it. Poor rotamers were left unchanged, only if additional interactions or tight packing and a clear electron density indicated a poor conformation. As mentioned in section 4.6.9, the lengths and angles of atomic bonds display only slight variations (Engh and Huber, 1991). Therefore distortions in length or angle of atomic bonds that are above 4σ are designated geometry outliers and often indicate poor local geometry.

The thermal displacement of atoms in protein structures is restrained by the covalent atomic bonds. Thus atoms closely linked should also have comparable *B*-factors. The rmsd of *B*-factors is, just as bond lengths and angles, restrained during refinement. It has therefore to be considered that adherence of the model to the ideal values is not an indicator of model quality, but rather an indicator of how strictly the restraints were enforced during refinement.

4.6.11 Structure Visualization and Analysis

Structure visualization was performed with PyMol. The structures in this work are mostly represented as cartoons (Richardson, 1985) with important amino acids represented as sticks. Small molecule ligands are usually displayed as sticks or balls and sticks. The assignment of secondary structure elements was done automatically with DSSP (Dictionary of Secondary Structure of Proteins) (Kabsch and Sander, 1983) by using H-bond, main chain torsion angle and distance criteria. For the calculation of solvent accessible protein surfaces, the Connolly method was used (Connolly et al., 1983). A sphere with radius 1.4 Å (representing a water molecule) is used to "probe" the surface of the molecule. Buried surface areas of protein complexes were calculated with PISA (Krissinel and Henrick, 2007) from the CCP4 package. Atomic distances were measured using COOT and PyMOL. Root mean square deviations were calculated using SUPERPOSE (Krissinel and Henrick, 2004). The analysis of protein ligand interactions was performed with MOE.

4.7 Molecular Dynamics Simulation

The simulation of the time dependent behavior of a system of atoms, such as a protein, is called molecular dynamics (MD). In this work, MD was used to gain insight into conformational rearrangements that have to take place between the transition

of the apo form of aDabi-Fab2 to its holo form complexed to dabigatran. For classical molecular mechanic approaches the protein is considered to be a conglomerate of balls corresponding to atoms with a fixed electronic distribution connected by springs that represent the atomic bonds (Nurisso et al., 2012). A force field that accounts for attracting and repelling forces e.g. coulomb forces, van der Waals forces, atomic bonds and H-bonds) is applied to the protein. The atoms in the protein get accelerated in certain directions as the force field acts on them. After each time step, the forces on the atoms are computed and combined with the current positions and velocities to generate new positions and velocities a short time ahead. The force acting on each atom is assumed to be constant during the time interval. The atoms are then moved to the new positions, an updated set of forces is computed and a new dynamics cycle starts (Schneider et al., 2008). The magnitude of these integration steps is limited by the fastest motions in the molecule. In proteins these are usually the bond vibrations between hydrogen and heavy atoms such as C, N, or O (Grenander, 1959). To allow for the calculations of longer time intervals, virtual sites can be used as explained later.

Prior to MD simulations with Gromacs version 4.5 (Hess et al., 2008), the solved X-ray structures of aDabi-Fab2 were edited with PyMol. The Fab was truncated to the FV part and water molecules that had a distance of more than 5 Å from the protein were removed. A topology file was created with pdb2gmx which adds hydrogen atoms to the model and contains all the information necessary to define the molecule within the simulation. This information includes nonbonded parameters (atom types and charges) as well as bonded parameters (bonds, angles, and dihedrals). The Amber99SB-ILND force field (Lindorff-Larsen et al., 2010) and the TIP3P water model (Jorgensen et al., 1983) were selected for subsequent simulations. A hexaedric simulation box with periodic boundary conditions was created with dimensions exceeding the FV by at least 1.2 Å in each direction. The box was populated with water as specified in the TIP3P solvent configuration. Sodium and chloride ions were added to compensate for the net charge of the protein with a final concentration of 150 mM. Prior to the productive MD run, the system consisting of the simulation box, the protein model, the water molecules and ions has to be relaxed through energy minimization. The resulting system represents a model with energetically favored geometry and solvent orientation, equilibrated at 300 K and a pressure of 1 bar. During the MD simulation, electrostatic interactions were calculated at every step with the particle-mesh Ewald method (Essmann et al., 1995). Short range repulsive and attractive dispersion interactions were described by a Lennard-Jones potential, with a 1.0 nm cut-off. The Settle algorithm (Miyamoto and Kollman, 1992) was used to restrain bond lengths and angles of water molecules and LINCS (Hess, 2008) was used

to restrain all other bond lengths. In addition, the fastest angular degrees of freedom involving hydrogen atoms were replaced with virtual interaction-sites (VSITE), allowing for a time step of 4 fs. The temperature was kept at 300 K through velocity rescaling (Bussi et al., 2008) interval (ô = 2.5 ps) and the pressure was controlled at 1 bar using the Parrinello-Rahman coupling scheme (ô = 5 ps) (Parrinello and Rahman, 1981). The trajectories were analyzed especially for the flexibility of CDR:H3. A principal component analysis was performed and the eigenvalues of the first two eigenvectors were plotted against each other and compared to the conformation of CDR:H3 observed in the apo and holo structures of aDabi-Fab2.

4.8 Biophysical Methods

4.8.1 Surface Plasmon Resonance

Surface plasmon resonance (SPR) is an electron charge density wave phenomenon that was first observed in the late 1950s (Turbadar, 1959). Surface plasmons are electromagnetic waves propagating along the interface between a thin metal layer and a dielectricum. The generation and usage of surface plasmons resonance for biomolecular interaction analysis will be explained on the basis of the Kretschmann-Raether configuration (Kretschmann and Raether, 1968) that is used in BIAcore T200 and most other SPR biosensors (see figure 8). P-polarized light is passed through a glass prism on which a thin gold film (∼50 nm) is evaporated. When a light beam propagates through two non-absorbing media with different refractive indices (e.g. glass and metal), it will be partially refracted and reflected at the boundary of these materials. Below a critical incident angle (Θ), the beam is no longer refracted but completely reflected, a phenomenon called total internal reflection. Upon reflection an evanescent electric field, that decays exponentially with distance, exists outside the metal and leeks into the adjacent dielectric medium containing the molecules to be analyzed (Boardman, 1982; Raether, 1997). At a certain angle of incident p-polarized light (SPR angle), the photons of the evanescent wave field excite electrons in the gold film, resulting in electron density fluctuations called surface plasmon waves. Although a complete description of this effect requires quantum physics, a simplified explanation of SPR is possible by describing the photons and plasmons involved as vector quantities. The photon momentum at the interface can be split into two vectors, one parallel and the second perpendicular to the interface. The relative magnitude of the components changes with the angle and wave-length of the incident light. In the same way, plasmons can be described as a vector, the magnitude of which depends on properties of the metal, thickness of the gold layer and the composition of the surrounding medium. When the parallel component of the

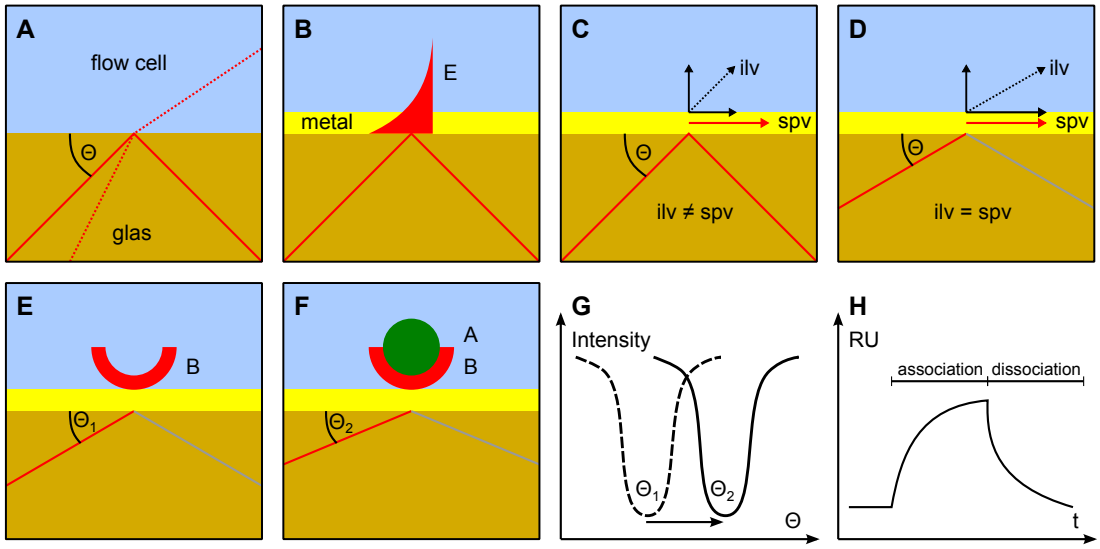


Figure 8: Basic Principles of Surface Plasmon Resonance. (A) P-polarized light gets either refracted (dashed line) or reflected (solid line) at the interface of media with different refraction index. Below a certain angle Θ, total internal reflection occurs. **(B)** The evanescent field E leaks through a metal film into the adjacent medium. It decays exponentially with increasing distance. **(C, D)** Only if the photon momentum of the incident light vector (ilv) parallel to the surface has the same magnitude as the surface plasmon vector (spv), photons will be converted to plasmons. This results in a sharp dip in intensity of the reflected light. **(E, F, G)** Binding of ligands to immobilized proteins changes the mass bound to the surface and results in altered resonance angles. **(H)** The sensogram shows the change of resonance angle in (RU) over time. A typical sensogram is composed of an association and a dissociation phase. k_a and k_d can be derived from the sensograms (adapted and modified from Ricklin, 2005).

incident light vector matches the surface plasmon, a resonance phenomenon occurs and photons are converted into plasmons. This results in a sharp dip in the intensity of reflected light at the SPR angle. The surface plasmon vector is influenced by the neighboring medium that contacts the gold film within a distance of 1 μm. In most SPR biosensors, a sample cell is directly attached to the gold film. Thus changes in the refracting index due to binding events in the vicinity of the gold layer immediately result in changes of the SPR angle that can be followed in real time (De Crescenzo et al., 2008).

In BIAcore biosensors the unit of the response signal is called resonance unit (RU) and represents a shift in resonance angle of approximately 10^{-4} °. It has been shown, for molecules with a high protein and low lipid and carbohydrate content, that there is a direct correlation between sensor signal and increase in mass on the sensor chip (Johnsson et al., 1991; Stenberg et al., 1991). The surface of the BIAcore chip contains a polysaccharide matrix of carboxymethylated dextran layers covalently attached to the gold layer on a small glass carrier. Dextran is a glucose-derived, long, unbranched and flexible polymer. The carboxymethylation creates a negative net charge that allows for covalent chemical linkage of proteins while, at

the same time, limits non-specific aggregation (Löfås and Johnsson, 1990). For a typical SPR experiment, one binding partner has to be immobilized while the other is injected in solution. Using amine coupling chemistry proteins can be covalently bound to the matrix via amine groups displayed on the N-terminus or on free lysines (Johnsson et al., 1991). When a binding partner to the immobilized protein is injected, the sensogram is dominated by the association phase (see figure 8). As formed complexes already dissociate in this phase, the SPR signal reaches a plateau when an equilibrium between association and dissociation arises. When injection of the binding partner is stopped and only buffer runs over the flow cell, the dissociation phase starts and the SPR signal steadily declines. The recorded sensogram contains valuable information, not only about the affinity of an interaction, but also about kinetic constants of association and dissociation. The simplest description of a binding event is a reversible 1:1 interaction model (Onell and Andersson, 2005):

$$A + B \rightleftharpoons AB \tag{39}$$

The association (K_A) and dissociation constant (K_D) are described by

$$K_D = \frac{[A] \cdot [B]}{[AB]} \tag{40}$$

$$K_A = \frac{[AB]}{[A] \cdot [B]} \tag{41}$$

where $[A]$, $[B]$ and $[AB]$ are the concentrations of the binding partners and complex, respectively. Equation (42) can be rewritten as a differential equation that describes the formation of AB over time

$$\frac{d[AB]}{dt} = k_a \cdot [A] \cdot [B] - k_d \cdot [AB] \tag{42}$$

where k_a [$M^{-1}s^{-1}$] is the rate of formation of new complexes and k_d [s^{-1}] is the rate of complex dissociation. At equilibrium, the change in concentration of AB is zero, leading to

$$0 = k_a \cdot [A] \cdot [B] - k_d \cdot [AB] \tag{43}$$

$$k_d \cdot [AB] = k_a \cdot [A] \cdot [B] \tag{44}$$

from equation (40) follows

$$\frac{k_d}{k_a} = \frac{[A] \cdot [B]}{[AB]} = K_D \tag{45}$$

typically, molecule *B* is immobilized on the chip and a known concentration of molecule *A* is injected over the surface. Equation (42) describes the progression of the interaction as well as equilibrium characteristics. When the supply of the binding partner in solution (i.e. molecule *A*) is too slow to maintain the same concentration near the surface of the sensor as in bulk, this is called mass transport limitation. Then equation (42) is no longer valid and needs to be expanded with a term describing diffusion (Karlsson and Fält, 1997; Myszka et al., 1998)

$$\frac{d[A]}{dt} = k_t \cdot ([A_{bulk}] - [A]) - k_a \cdot [A] \cdot [B] + k_d \cdot [AB] \tag{46}$$

$$\frac{d[B]}{dt} = -k_a \cdot [A] \cdot [B] + k_d \cdot [AB] \tag{47}$$

$$\frac{d[AB]}{dt} = k_a \cdot [A] \cdot [B] - k_d \cdot [AB] \tag{48}$$

with $[A]_{t=0} = 0$, $[B]_{t=0} = R_{max}$ and $[AB]_{t=0} = 0$. In equation (46), $[A]$ is the actual concentration of molecule *A* at the sensor surface and $[A_{bulk}]$ is the known concentration of molecule *A* in solution. The mass transport coefficient k_t [RU M^{-1} s^{-1}] describes the transport rate of *A* to the sensor surface and R_{max} [RU] is the maximum binding capacity of the sensor. Usually a concentration series of molecule *A* is injected over the chip surface, resulting in multiple sensograms. To these sensograms the values of k_a, k_d, k_t and R_{max} are globally fitted with an optimizer and, as equations (46 - 48) cannot be solved analytically, a numerical integrator. The affinity K_D can be computed from the obtained kinetic rates with equation (45). The measure for the accuracy of the fitted parameters χ^2 describes the residual noise per data point after fit, according to

$$\chi^2 = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n - p} \tag{49}$$

where $\hat{y}_i$ is the estimate of the measured data point y_i, *n* is the number of data points and *p* is the number of parameters fitted by the optimizer. Besides the 1:1 interaction

model, there are more complex models that take into account possible conformational changes of the binding partners or inhomogeneous protein preparations. With additional parameters, these models are capable of fitting more complex sensograms but a good fit to the data is no evidence that the proposed mechanism is correct. So use of these models is generally not recommended unless there is strong biological evidence for a deviation of the standard 1:1 model (Rich and Myszka, 2002).

4.8.2 Isothermal Titration Calorimetry

Almost any chemical reaction or physical change in a system, such as a binding event, is accompanied by a change in heat or enthalpy. When heat is taken up from the surroundings the process is called endothermic, when heat is released the process is referred to as exothermic. The change of enthalpy at a constant temperature can be measured by the technique of isothermal titration calorimetry (ITC). In a single experiment, the values of affinity (K_A and K_D, respectively), the stochiometry (n) and the binding enthalpy (ΔH_b) can be determined. The free energy and entropy of binding are calculated from the association constant (Pierce et al., 1999). The basic features of an ITC power compensation instrument and an exemplary output of raw data is shown in figure 9 (Ladbury and Doyle, 2004). Two identical cells, one for the binding reaction to take place (the sample cell) and one to act as a reference, are located in an adiabatic jacket. The adiabatic jacket is cooled by a circulating water bath to a temperature below that at which the experiment is conducted. Both cells of the calorimeter are controlled at constant temperature (thus isothermal). The sample cell contains one of the binding partners (component *A*) and a stirring syringe which contains component *B*. The reference cell contains a solution of similar heat capacity as in the sample cell (usually buffer or water). Heaters located on both cells are employed to maintain identical temperatures between the cells. Sensitive thermocouple devices are used to monitor any heat change upon stepwise titration of component *B* to component *A*. If the binding reaction is exothermic, the sample cell will require less power to maintain thermal equilibrium with the reference cell. This change in power input is directly measured and plotted in a thermogram. The heat change upon each injection can be calculated by simply integrating the heater power over the time of the measurement (Freyer and Lewis, 2008).

The standard thermodynamic relationship is described as

$$\Delta G = \Delta G^{\circ} + RT \ln K_A \tag{50}$$

where *T* is the absolute temperature in Kelvin, *R* is the gas constant,ΔG is the Gibbs

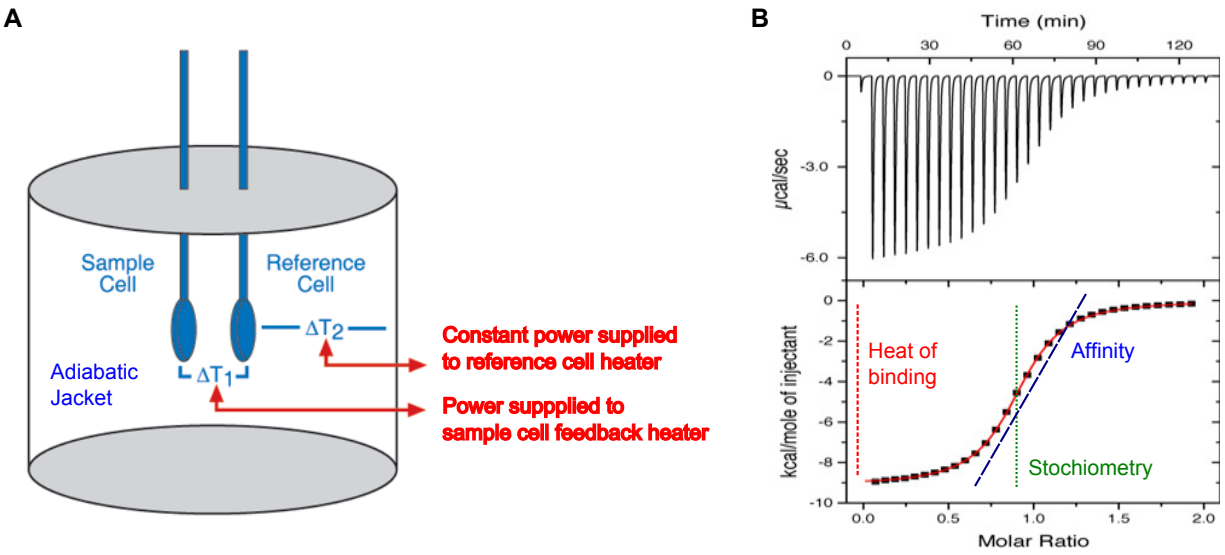


Figure 9: Basic Principles of Isothermal Titration Calorimetry. **(A)** The calorimeter is composed of an adiabatic jacket that harbors two identical cells. Constant heat power is supplied to the reference cell, the power necessary to keep the sample cell at constant temperature upon injection of the binding partner is recorded. **(B)** The exemplary power signal of the raw file (upper half) is integrated for each injection to obtain the thermogram (lower half). Heat of binding, affinity and stochiometry can be deduced from the thermogram as indicated (adapted and modified from Freyer and Lewis, 2008).

free energy of the process under study, $\Delta G°$ the standard Gibbs free energy and K_A is the reciprocal of the dissociation constant K_D. Equation (50) can also be written as

$$\Delta G = \Delta H - T\Delta S \tag{51}$$

where ΔH is the enthalpy change, T the absolute temperature [K] and ΔS is the change in total entropy of the system. Reactions with similar affinities can have different enthalpic and entropic contributions to their free energies. The equilibrium and mass balance relationships of a 1:1 binding experiment can be described with equations (52) and (53), respectively.

$$\Theta_j = \frac{[L] \cdot K_A}{1 + [L] \cdot K_A} \tag{52}$$

$$L_t = [L] + P_t(n \cdot \Theta) \tag{53}$$

where Θ_j is the fraction of site *j* occupied by ligand, L_t is the total ligand concentration, $[L]$ is the free ligand concentration, P_t is the total macromolecule concentration, K_A is the binding constant of the reaction and *n* is the stochiometric ratio. Substi-

tution of $[L]$ into equation (52) allows for the calculation of the fraction of binding sites Q that are occupied.

$$Q = P_t V_0 (n\Theta\Delta H) \tag{54}$$

$$\Delta Q(i) = Q(i) - Q(i-1) \tag{55}$$

The total heat produced can be calculated from equation (54), where V_0 is the initial volume of the sample cell and ΔH is the molar enthalpy change. The differential heat is defined by equation (55), where i represents the injection number. The heat changes in the raw thermogram are integrated and corrected for dilution effects by subtraction of the appropriate blank experiments with the Origin 7 ITC software package (MicroCal). Nonlinear regression is then performed on the parameters K_A, n and ΔH to obtain the best fit to the measured data. ΔS is subsequently calculated from the three parameters fitted.

An iTC200 (MicroCal) was used to determine the thermodynamic properties of the binding of dabigatran to thrombin and aDabi-Fabs. A weighted amount of crystalline dabigatran was solubilized in ITC buffer. The proteins were dialyzed against the same buffer. The experiment was run at 25 °C and was set to deliver 20 (2 μl) injections at 150 sec intervals. Dabigatran (100 μM) was gradually injected into 330 ml of protein solution (10 μM). To correct for dilution effects, dabigatran (100 μM) was also injected to the sample cell containing only buffer but no protein components in a subsequent experiment. Thrombin as well as the Fabs possess only one specific binding site for dabigatran, thus a 1:1 binding model was assumed for analysis of the ITC data.

4.8.3 Kinetic Exclusion Assay

The KinExA 3200 (derived from Kinetic Exclusion Assay) (Sapidyne Instruments) is a liquid handling system equipped with a fluorescence detector that can quantify the concentration of a molecule in solution and can be used to determine the affinity of a binding reaction (see figure 10 on the facing page). The concentration is determined by using a small column of beads which contain one partner immobilized that specifically binds the molecule to be analyzed (Darling and Brault, 2004). The two molecules *A* and *B* are mixed and incubated to achieve complex formation. In case of a 1:1 interaction, a mixture of three species *A*, *B* and the complex *AB*, exist. The KinExA assay allows for determination of the concentration of one of the unbound components (e.g. *B*). Therefore a small column of polymethylmethacrylate (PMMA)

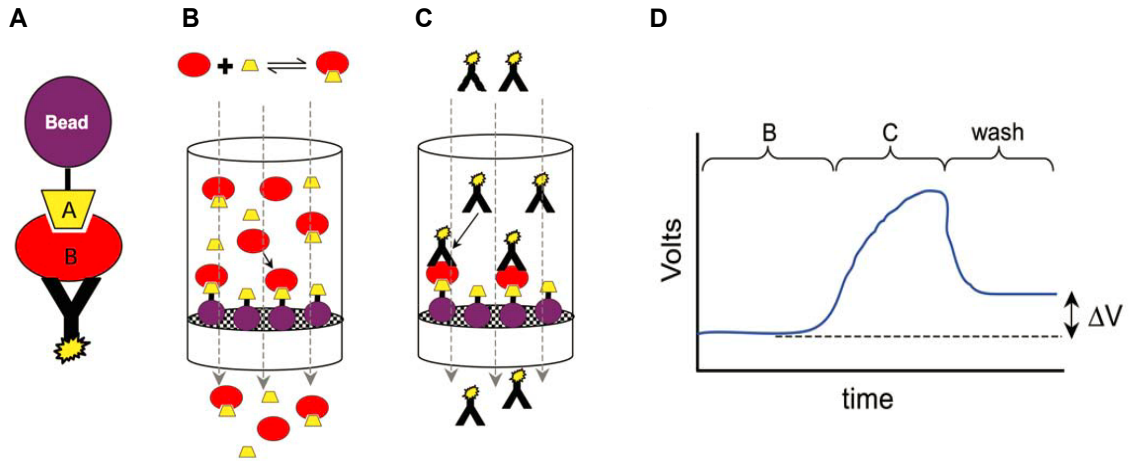


Figure 10: Principle of Kinetic Exclusion Assay. (A) Molecule A is immobilized to polystyrene beads. B free in solution binds to immobilized A and a fluorescence labeled antibody against B is used to quantify the amount of B captured on the beads. **(B)** The beads are transferred to a micro column and the mixture of A, B and AB is allowed to flow through for a short period of time. Immobilized A captures a fraction of B that is proportional to the amount of free B in solution. **(C)** A fluorescent antibody against B is applied to detect the amount of B captured. **(D)** After rinsing away excess fluorescent antibody, the change in fluorescence (ΔV) is determined. Adapted and modified from Darling and Brault (2004).

beads with *A* (or a close analog) immobilized is packed by the instrument and the mixture containing *A*, *B* and *AB* is applied to that column. As long a the amount of immobilized *A* is in excess compared to *A* in solution, the amount of *B* captured is proportional to the concentration of *B* in solution (Blake et al., 1999). When the contact time between mixture and beads is short enough, significant dissociation of *AB* in solution as *B* is partially depleted is “kinetically excluded”. A fluorescence labeled antibody that binds to *B* is then passed through the beads and the fluorescence emission is measured in volts.

In KinExA experiments the total concentration of each component $[A_0]$ and $[B_0]$ is known. From the law of conservation of mass it follows that

$$[A_0] = [A] + [AB] \tag{56}$$

$$[B_0] = [B] + [AB] \tag{57}$$

When equations (56) and (57) are included into equation (45), the resulting quadratic equation can be solved for $[B]$ which is proportional to the instrument signal

$$Signal = \frac{Sig_{100\%} - Sig_{0\%}}{2[B_0]} \left(([B_0] - K_D - [A_0]) \right.$$

$$+\sqrt{[B_0]^2 + 2[B_0]K_D - 2[B_0][A_0] + K_D^2 + 2[A_0]K_D + [A_0]^2}\Big) + Sig_{0\%} \quad (58)$$

where $Sig_{100\%}$ is the voltage signal produced by B_0 in the absence of A, and $Sig_{0\%}$ is the voltage signal in the absence of B. The optimal solution for $[B_0]$, K_D, $Sig_{100\%}$ and $Sig_{0\%}$ (A_0 is known) is determined by minimizing the squared error between the measured data points and equation (58). The affinity of dabigatran (and close analogs) for aDabi-Fab1 was measured using KinExA 3200 technology. As described above, one of the binding partners must be immobilized on a solid phase. For this purpose, neutravidin was coupled to beads and subsequently used to capture biotinylated dabigatran. Neutravidin is a tetrameric protein with an exceptional strong affinity for biotin ($K_D = 10^{-15}$ M) that is often used for permanent capture of biological molecules in a variety of different assays (Wilchek et al., 1988; Kobayashi et al., 1995). 1 ml of Neutravidin (1 mg/ml) (Sigma-Aldrich) was immobilized on 0.5 ml of PMMA beads (Sapidyne Instruments) by incubation for 8 h at 4 °C on a roller. Beads were centrifuged for 5 min at 1000 x g and the supernatant discarded. Beads were resuspended and washed with 1 ml PBS. 0.5 ml of biotinylated dabigatran (0.1 mg/ml) was added and incubated for 15 h at 4 °C. Beads were washed 2 times with PBS and blocked by incubation with 1 ml of BSA (10 mg/ml) (bovine serum albumin). For accurate determination of the affinity, the concentrations used during measurement should not be larger than 10 times the value of K_D. aDabi-Fab1 was used in concentrations of 25 - 100 pM. To prevent unspecific absorption on tube walls while preparing the solutions or later in the instrument, all following steps were performed in PBS supplemented with 1% BSA. For each compound a 2-fold dilution series was prepared and incubated with a constant concentration of aDabi-Fab1 for 24 h at 4 °C (for detailed information see table 8 on the next page). The solutions were injected over the freshly packed micro column for 528 s at a flow rate of 0.25 ml/min. Captured aDabi-Fab was detected by injection of 500 ml of a Cy5-labeled anti-human light chain Fab from goat (Rockland Immunochemicals) with a final concentration of 13 pg/ml by measuring the fluorescence at 650 nm.

Table 8: Concentrations for Affinity Determination with KinExA

compound	c_{max}	c_{min}	c(aDabi-Fab1) [pM]
dabigatran			
1	64 µM	62.5 nM	100
2	32 µM	31.25 pM	100
3	521 pM	250 fM	25
4	1 nM	16.93 fM	25
5	32 nM	31.25 pM	25
6	256 pM	250 fM	25
7	32 nM	31.25 pM	25
8	64 pM	62.5 fM	25
9	32 nM	31.25 pM	100
10	64 pM	62.5 fM	25
11	64 pM	62.5 fM	25
12	n.d.	n.d.	nb
13	3.2 µM	3.13 nM	100
benzamidine	64 µM	62.5 nM	100

5 Results and Discussion - Antibody Fragments Against Dabigatran

Dabigatran is a direct thrombin inhibitor that is used for anticoagulation therapy such as stroke prevention in patients with atrial fibrillation. As every anticoagulation therapy bears the risk of bleeding events, antibody fragments (Fabs) against dabigatran were developed. These Fabs may be used for the rapid reversal of dabigatran's anticoagulative effects in emergency situations. The X-ray structures of these Fabs in complex with dabigatran were solved and are described thoroughly during this section. Fab1 binds to dabigatran with extremely high affinity and reverses its effects in a rat model in vivo. The high affinity is achieved by the structural mimicry of interactions with dabigatran that are also formed by thrombin. Fab2 shows lesser similarities with thrombin but also a lower affinity for dabigatran. There are large conformational rearrangements between the apo and complexed form of the Fab. CDR:H3 shields a hydrophobic area at the binding site and opens in the presence of dabigatran to allow for binding.

5.1 aDabi-Fab1

5.1.1 Antibody Screening and Engineering

The generation of dabigatran specific monoclonal antibodies was performed by the NBE unit. For this purpose, mice were immunized with dabigatran derived haptens

(figure 4, compound 3) coupled to carrier proteins. Isolated antibodies were screened for binding to the immunogen and for specific binding to soluble dabigatran in a competitive ELISA. Antibodies showed binding to a hapten-peroxidase conjugate and could be competed off by increasing concentrations of dabigatran. For the purpose of therapeutic application in humans it is necessary to minimize the immunogenicity of proteins. Therefore the antibodies were engineered by grafting the complementarity-determining regions (CDRs) onto human Ig variable region sequences and a library of various humanized Fab molecules were screened for binding to dabigatran. The functional inhibition of dabigatran activity was tested in a modified thrombin time assay. Clotting time was prolonged from 25 to 100 s in the presence of 7 nM dabigatran. When the dabigatran-specific antibodies were added in increasing concentrations, the anticoagulant activity of dabigatran was reversed in a concentration-dependent manner. The parental mouse antibodies, mouse Fab fragment and humanized aDabi-Fab1 all showed similar inhibitory activity with IC50s between 2 to 4 nM. Clotting times were reduced to baseline at antibody concentrations above 10 nM, indicating complete neutralization of dabigatran. The antibody alone in the absence of dabigatran had no effect on the coagulation time. Similarly, a Fab not specific for dabigatran, did not reverse the anticoagulant effect of dabigatran. The affinities of mouse and humanized Fabs for dabigatran were measured in solution. While the parental mouse Fab had a K_D of 41 pM, the humanized Fab displayed further increased binding with a K_D of 2.1 pM. The Fabs displayed similar binding affinities as full-length antibodies which shows that the Fc region does not influence dabigatran neutralization potency. The high affinity of aDabi-Fab1 correlates with an extremely slow off-rate of $\sim 0.7 \times 10^{-6}$ s, almost 200-fold lower compared to the parental mouse antibody. In order to better understand the molecular mechanism of the dabigatran specific Fab, structural studies were performed.

5.1.2 Crystallization of Fab1

aDabi-Fab1 was obtained from the Biopharmaceuticals department of Boehringer Ingelheim and, prior to crystallization, purified by SEC on HiLoad-26/60-Superdex-75 prep grade (GE Healthcare Life Sciences) equilibrated and run in standard protein buffer. The complex of aDabi-Fab1 with dabigatran (or analogs, see section 5.2) was formed by incubation of the protein with 2.5 molar excess of the small molecule for 1 h at 4 °C. Prior to crystallization, the complexes were concentrated to 10 mg/ml and the buffer was changed to water. Crystallization trials were set up using the vapor diffusion method by both the hanging and the sitting drop procedures. 1 μl of the complex was mixed with an equal volume of 15 mM di-sodium hydrogen phosphate

(pH 11) and 24% PEG 3350 and equilibrated against 24% PEG 3350. 1 μl of uncomplexed Fab1 was mixed with an equal volume of 25 mM tri-sodium citrate (pH 3) and 2% PEG 3350 and equilibrated against 24% PEG 3350. Crystals appeared after 5 days of incubation at 20 °C

5.1.3 Overall Structure of the Complex

Detailed information on crystallography is given in appendix A for all structures presented in this work. The crystal structure of the antigen binding fragment aDabi-Fab1 in complex with dabigatran was solved at 1.7 Å resolution. Two crystallographically independent Fab:dabigatran complexes are located in each asymmetric unit of the crystal. For convenience the complex consisting of Fab chains H + L and A + B will be named Fab 1 and Fab 2, respectively. As immunoglobulins all share a common fold, the structural features will be discussed thoroughly only for the aDabi-Fabs (for the antibodies against PCSK9, a more general discussion focused rather on their modes of action will be given in section 6). aDabi-Fab1 is composed of two variable domains ($V_L\varkappa$ and V_H) and two constant domains ($C_L\varkappa$ and C_HIgG1). These domains show the typical immunoglobulin fold (Bork et al., 1994). The antigen binding site of antibodies is formed by the complementarity-determining regions (CDRs) on the surface of the variable domains which are named from L1 to H3. The structures of the two Fab:dabigatran complexes overall are very similar. The observed root-mean-square distance (rmsd) over all Cα-atoms of 1.1 Å is due to differences in the relative subunit orientations and conformational differences in loop regions of the paratope. The angle between the two pseudo-dyad axes relating the two variable and constant domains is called the elbow angle. Fab1 displays an elbow angle of 135.1°, Fab 2 of 143.2°. There are two significant structural differences between the protomers. The conformation of the heavy chain C-terminal region from H:Glu212 to H:Cys216 (numbering according to Wu and Kabat (Johnson and Te Wu, 2001)) is different because of crystal packing. Moreover in CDR:H3 the tyrosines H:Tyr100 and H:Tyr100A adopt different side chain conformations, also due to crystal contacts. The electron density is well defined for the major part of the model all the way to the cysteines that form the intermolecular disulfide bond. Only for Fab 2 there are two ill-defined stretches, that is from L:Ser26 – L:Leu27B in CDR:L1 and from H:Ser127 – H:Ser132 in the constant domain of the heavy chain. Dabigatran has the same conformation in both protomers with a rmsd of 0.52 Å over all atoms. Usually the CDR loops adopt conformations that can be grouped into canonical structural classes (Al-Lazikani et al., 1997). CDR:L1 ranges from residue 24 – 34 of the light chain and has a 5-residue insertion at position 27. It belongs to canonical structure

class 4. The residues 50 – 56 build CDR:L2 that resembles class 1. Like in all other antibody structures published, the amino acid on the tip of this loop (L:Val51) adopts a strained conformation and contributes to the formation of a γ-turn. The third CDR in the light chain (residues 89 – 97) has a conformation typical for canonical structure 1. CDR:H1 spans residues 26 – 35 and is similar to the canonical structure class 1. The CDR:H2 loop adopts canonical class 1. For CDR:H3 that displays an insertion of 6 amino acids at position 100, no classical canonical structure can be assigned. The region of CDR:H3 proximal to the framework is called the torso. Depending on the conformation, the torso can be classified as bulged or non-bulged (Morea et al., 1998). aDabi-Fab1 adopts a bulged conformation the characteristic structural feature of which is the hydrogen bond (H-bond) between the side chain nitrogen of H:Trp103 and the backbone carbonyl oxygen of the residue 3 positions N-terminal to it, in this case H:Phe100F. The β-bulge is stabilized by an interaction of the carbonyl oxygen of H:Ser94 that acts as H-bond acceptor for the backbone nitrogen atoms of H:Ala101 and H:Tyr102, respectively. The light chain further contributes to stabilize the bulge. L:Arg46 forms an H-bond to the carbonyl oxygen of H:Gly100E and salt bridges to H:Asp100E and L:Asp55 which is located in CDR:L2.

5.1.4 Crystal Structure Explains High Affinity

The dabigatran binding site of the Fab (The further description of structural features refers to Fab 1 as this is the protomer which is completely defined by electron density. If the respective values differ significantly for Fab 2, this will be mentioned explicitly.) is located at the interface of the variable domains (see figure 11 on the facing page). The strong affinity of 2.1 pM of aDabi-Fab1 for dabigatran can be qualitatively explained by the tight interactions that are revealed by the costructure. All CDR loops except L2 are involved in dabigatran binding. The binding is mediated by hydrophobic interactions, H-bonds and a salt bridge. The benzamidine moiety of dabigatran inserts into a cavity formed by heavy and light chain, whereas the benzimidazole, carboxamide and pyridine moieties are partially exposed to solvent. The propionic acid moiety of the hapten is almost completely exposed to solvent in Fab 2, only contacting the hydroxyl group of H:Ser56. In Fab 1, the carboxylate oxygens form H-bonds to H:Gly133 nitrogen and the H:Thr135 sidechain oxygen of a symmetry mate, respectively. The fact that the carboxylate moiety is not involved in binding to the Fab is consistent with the employment of the hapten conjugated at this position to a carrier protein for immunization. The interface area (calculated as difference in total accessible surface areas of isolated and interfacing structures divided by two) of antibody and dabigatran is 531 Å^2 (505 Å^2 in Fab 2) To this number the light chain

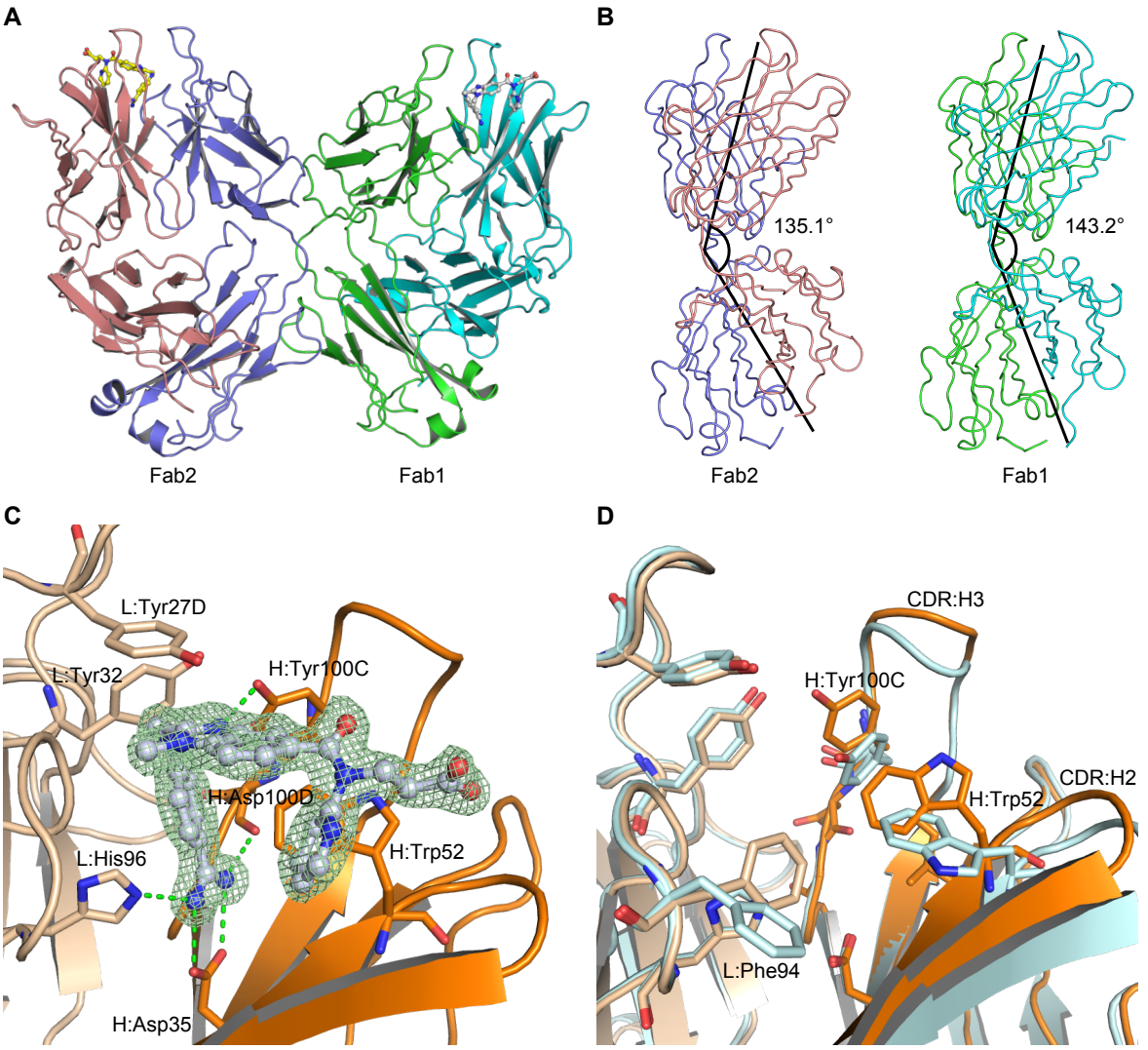


Figure 11: Structure of aDabi-Fab1. (A) Two protomers are present in the asymmetric unit. Their overall structure is very similar with deviations in those parts that are involved in crystal contacts and in the elbow angle. **(B)** The elbow angles for the two protomers are illustrated. **(C)** Binding site of aDabi-Fab in complex with dabigatran (the respective electron density of dabigatran is shown in pale cyan, contoured at 1.5 sigma). The benzamidine group of dabigatran extends into a cavity formed by the interface of light (light ocre) and heavy chain (dark ocre). It forms a bidentate salt bridge to H:Asp35 and additional H-bonds to L:His96 and H:Asp100D, respectively (indicated as green dotted lines). The amine nitrogen of dabigatran forms a nonclassical H-bond to L:Tyr32. L:Tyr27D interacts with the benzimidazole moiety of dabigatran via parallel displaced π-stacking. H:Tyr100C forms a H-bond to the aldimine nitrogen. H:Trp52 forms two T-shaped aromatic interactions with the benzamidine and pyridine moiety of dabigatran, respectively. **(D)** Superposition of the binding site of aDabi-Fab in its apo conformation (colored in pale cyan) and bound to dabigatran. The orientation of the sidechain of H:Tyr100C changes by 3.2 Å (hydroxyl oxygen) upon binding of dabigatran which causes the apex of CDR:H3 to increase its relative distance to the binding site by 2.1 Å (measured from the Cα of H:Tyr100). About the same distance CDR:H2 moves towards the binding site (measured from the Cα of H:Gly55) and the H:Trp52 and L:Phe94 adopt different side chain conformations. This concerted movement opens the binding site of dabigatran.

contributes 218 Å^2 (219 Å^2), the heavy chain 323 Å^2 (289 Å^2).

Hydrophilic interactions The most prominent interaction is conveyed via the benzamidine moiety of dabigatran. It extends deeply into a cavity that lies at the interface of heavy and light chain. The amidine group of dabigatran forms a bidentate salt bridge to H:Asp35 in CDR:H1 with nitrogen-oxygen distances of 2.9 Å and 3.0 Å, respectively. This interaction is additionally stabilized: one of the nitrogens forms a H-bond to Nε of L:His96 (2.9 Å) in CDR:L3, the other acts as a H-donor for the backbone carbonyl oxygen of H:Asp100D (3.2 Å) in CDR:H3. These polar interactions are expected to contribute strongly to the free energy of binding as they reside in an otherwise hydrophobic environment provided by the side chains of L:Phe94, H:Ile33, H:Val50 and H:Phe100F, which provide efficient solvent shielding. An additional H-bond is formed between H:Tyr100C, located in CDR:H3, and the aldimine nitrogen of the benzimidazole moiety of dabigatran. Here the hydroxyl group of H:Tyr100C acts as H-donor with an oxygen-nitrogen distance of 2.65 Å and a NHO-angle of 160°, indicating a fairly strong H-bond (Jeffrey and Saenger, 1994).

Aromatic interactions In protein structures and protein ligand interactions, two general orientations are observed between interacting aromatic rings. The T-shaped "edge-to-face" and the (often parallel displaced) stacked geometry (Burley and Petsko, 1985). As the parallel displaced geometry is observed more frequently, it is thought that it is energetically favored over the T-shaped (McGaughey et al., 1998). Both geometries are observed in the binding of aDabi-Fab1 to dabigatran. H:Trp52 in CDR:H2 contributes to the binding via two T-shaped π-interactions with the benzamidine ring (3.7 Å closest distance of rings) and the pyridine ring of the hapten (3.5 Å), respectively. Two tyrosines of CDR:L1 also interact with the hapten. L:Tyr27D is involved in parallel displaced π-stacking with the benzimidazole moiety with a distance of 3.4 Å between L:Tyr32 and the imidazole ring. L:Tyr32 displays a N-H/π interaction with the amine nitrogen of dabigatran. In this nonclassical H-bond the π-electrons of the phenyl ring act as H-bond acceptor (Burley and Petsko, 1986).

5.1.5 Apo Structure of aDabi-Fab1

The crystal structure of the uncomplexed aDabi-Fab1 was solved at 1.89 Å resolution with one protomer in the asymmetric unit with an elbow angle of 137.4°. With the exception of the sidechains of H:Tyr100 and H:Tyr100A in CDR:H3 the electron density is completely defined, even for those parts that are not defined in the complex structure. The CDR loops adopt similar conformations as observed in the complex

structure. Noteworthy, the binding site for dabigatran is occupied by two glycerol molecules originating from the cryo buffer. The first glycerol is located at the bottom of the binding site and forms H-bonds with H:Asp35 (oxygen-oxygen distance of 2.6 and 3.0 Å) and with the Nε-H of L:His96 (2.8 Å). The second glycerol forms H-bonds with the first glycerol (2.7 and 3.1 Å, respectively) and with H:Tyr100C (2.9 Å), L:Ser91 (2.7 Å), L:Thr92 (2.7 and 3.0 Å) and interacts with the aryl groups of L:Tyr27D and L:Tyr32.

5.1.6 Structural Differences Between the Protomers

The two protomers in the asymmetric unit of the crystal structure of aDabi-Fab1 in complex with dabigatran display elbow angles of 135.1° and 143.2°, respectively (Stanfield et al., 2006). Both values are in good agreement with the median of 150° for elbow angles of kappa chains found in the pdb most frequently. The fact that the same Fab displays elbow angles differing by 8° in the same crystal illustrates the general flexibility of the V-C interfaces and seems to be due to crystal packing (Stanfield et al., 1993). In addition to the elbow angles there are more conformational differences between the two protomers, especially for the antigen binding site. In the conformation seen in Fab 1, H:Tyr100 is oriented towards the hapten and forms a H-bond to H:Thr135 of a symmetry mate. In Fab 2, both tyrosines side chains are facing away from the hapten and H:Tyr100 forms a lattice contact with the amide nitrogen of H:Gly10 of a symmetry mate of Fab 1. Although the differing conformations of H:Tyr100 and H:Tyr100A appear to be due to crystal packing, together with the poor definition in the apo structure, it indicates that these tyrosines have an intrinsic flexibility and do not strongly contribute to the binding of dabigatran. More important for the binding is that the benzamidine moiety of dabigatran binds deeply into the combining site of the antibody and interacts there with H:Asp35. The tendency to form a concave paratope has been reported for antibodies recognizing small antigens and haptens (MacCallum et al., 1996). The difference in the interface area between the two Fabs is also mainly due to the conformation of H:Tyr100 and H:Tyr100A. In Fab 1 531 Å^2 of surface area are buried, in Fab 2 505 Å^2, respectively. In Fab 1 the tyrosines face towards dabigatran. Especially for H:Tyr100, this results in a decrease of surface area upon binding that is not observed in Fab 2. But with a distance of 4.39 Å between the closest non-hydrogen atoms (hydroxyl oxygen of H:Tyr100 and the carboxamide oxygen of dabigatran), this residue does not contribute to binding. The high rmsd of 1.10 Å over all Cα-atoms can be explained by the differences in the elbow angles and the conformations of the C-terminal region of the heavy chain and of H:Tyr100 and H:Tyr100A, all due to crystal packing. The light chain displays

a rmsd of 0.67 Å, but when the $V_L\varkappa$ and $C_L\varkappa$ domain are analyzed separately the rmsd decreases to 0.19 Å and 0.26 Å, respectively. This indicates that the flexibility of the V-C interface mainly accounts for the high rmsd of the light chain. The rmsd of the heavy chain is 1,27 Å. For the V_H and the C_HIgG1 domain the rmsd is 0.58 Å and 0.67 Å, respectively. The larger rmsd compared to the light chain can be explained by the different conformations of the C-terminal region and of H:Tyr100 and H:Tyr100A.

5.1.7 Conformational Changes Between Apo and Holo Structure

The comparison of the free Fab with the dabigatran complex reveals no major conformational change of the protein backbone at the binding site induced by the hapten, but some major side chain reorientations (see figure 11). The two glycerol molecules seen in the binding site of the apo structure originate from the cryo buffer, thus one can expect the binding site to be occupied by water molecules under physiological conditions. Only two CDRs in the heavy chain show a small concerted movement upon binding: the orientation of the sidechain of H:Tyr100C changes by 3.2 Å (hydroxyl oxygen) upon binding of dabigatran which causes the apex of CDR:H3 to increase its relative distance to the binding site by 2.1 Å (measured from the Cα of H:Tyr100). CDR:H2 moves about the same distance towards the binding site (measured from the Cα of H:Gly55) and the side chain conformations of H:Trp52 and L:Phe94 are changed. This concerted movement allows for the binding of dabigatran as it provides sterical space that is otherwise occupied by the sidechains of H:Trp52 and H:Tyr100C in the apo structure. In the apo structure H:Trp52 is involved in a crystal contact, but this sidechain conformation seems to be a low energy conformation in the absence of dabigatran as it was also observed in a co-structure of aDabi-Fab1 with benzamidine that crystallized in a different spacegroup. A MD simulation was performed to further investigate the flexibility of the binding interface. Upon removal of dabigatran from the Fab its dynamics were simulated for 50 ns. As already inferred from analysis of the apo and holo structures, CDR:H3 highly flexible during the simulations. Interestingly for H:Trp52 both apo and bound conformations of the were observed. Although the holo conformation was predominant during the MD simulation, it is sound conclude that both are prevalent conformers in solution. This indicates that binding of dabigatran to aDabi-Fab1 is not based on an induced fit mechanism but rather on conformational selection of a preexisting sub state of the Fab by dabigatran. The same binding mode has also been shown for the binding of substrates to thrombin (Niu et al., 2011)

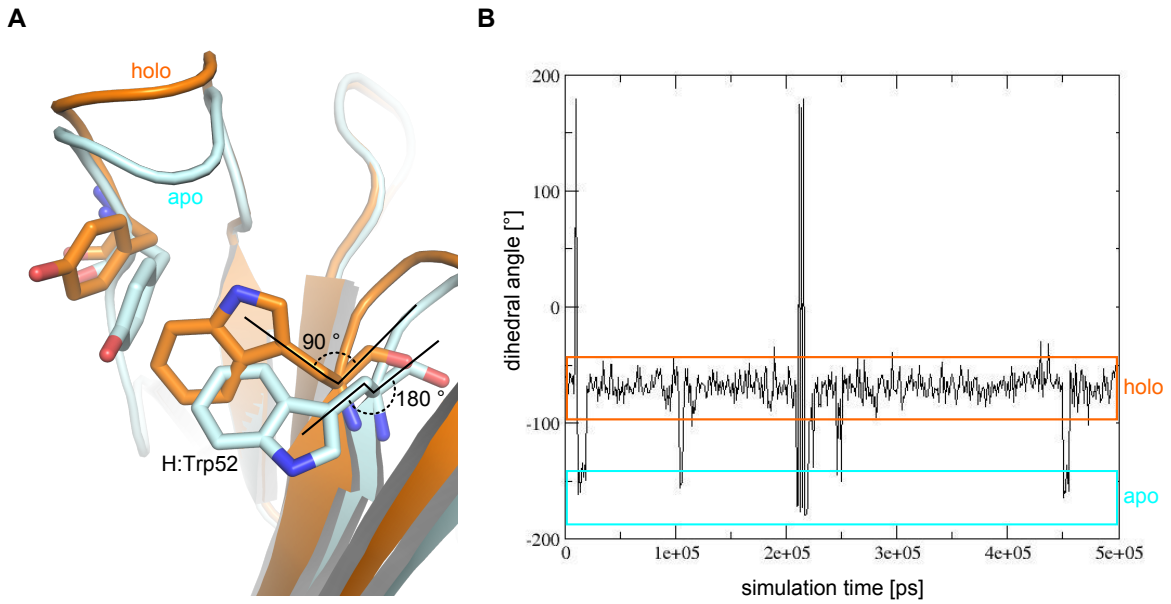


Figure 12: Molecular Dynamics Simulation of aDabi-Fab1. (A) Changes in the dihedral side chain angle of H:Trp52 were analyzed during a MD simulation. The crystal structures of aDabi-Fab1 in its apo and holo conformation are shown. H:Trp52 and H:Tyr100A are depicted in sticks as these are the residues with the larges conformational differences. **(B)** The dihedral angle as depicted in (A) during a MD simulation over 50 ns is shown. H:Trp52 rapidly switches between the conformations seen in the apo and holo structures. Interestingly both conformations of H:trp52 (with the holo conformation much more populated) are observed, indicating a conformational selection mechanism for the binding of dabigatran.

5.1.8 Fab Mimics Structural Features of Thrombin

The first crystal structure of thrombin was reported in 1989 by Bode et al. (Bode et al., 1989). In 1992 the structure of bovine thrombin in complex with a benzamidine based peptidic inhibitor NAPAP was published (Brandstetter et al., 1992). This structure served as starting point for a structure-based effort leading to discovery of dabigatran (Nar, 2012). During this process, costructures of thrombin in complex with several compounds were solved and deposited in the PDB (see PDB:IDs 1KTS, 1KTT, 1G30, 1G32) including the co-structure of thrombin and a dabigatran ethylester (Hauel et al., 2002; Nar, 2012). When comparing the structure of the thrombin:dabigatran complex with the costructure of aDabi-Fab1, it becomes obvious that dabigatran adopts different bound conformations. When superimposing the benzimidazole scaffold of dabigatran, the carboxamide and carboxylate group as well as the pyridine ring are very similarly positioned (rmsd 0.35 Å), while the benzamidine moiety is twisted by 66° due to a difference in torsion angles around the exocyclic bond at the benzimidazole. Despite the distinct bound conformations, the way thrombin and aDabi-Fab1 bind dabigatran are strikingly similar (see figure 13 on the next page). Most notably, a salt bridge anchors the benzamidine moiety of

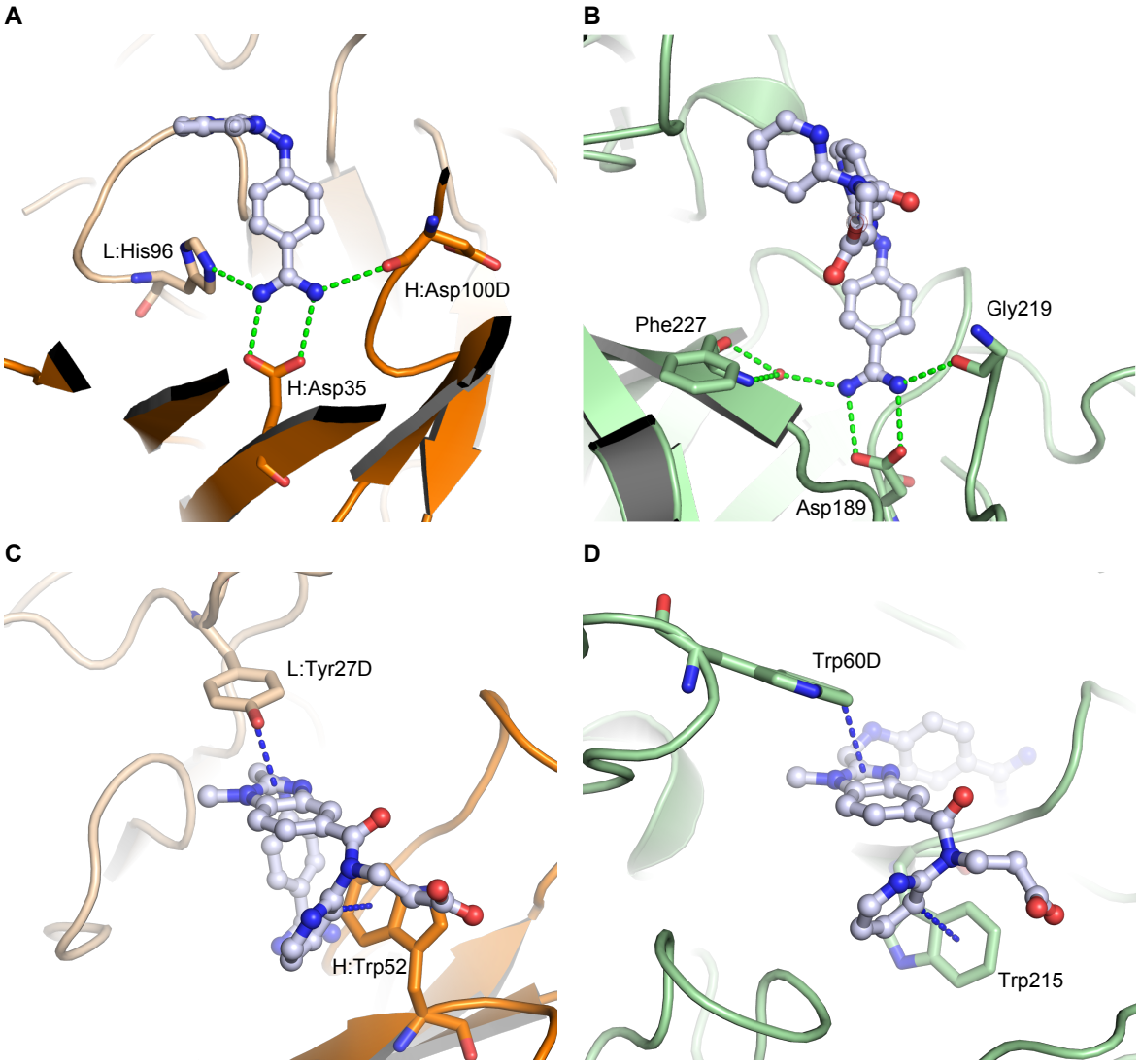


Figure 13: aDabi-Fab1 Mimics Structural Features of Thrombin. The dabigatran complex structure of aDabi-Fab and thrombin is displayed in the left and right panels, respectively (Fab is colored in orange, thrombin is colored in pale green, H-bonds and aromatic interactions are indicated as dotted lines in green and blue). **(A-B)** The benzamidine moiety of dabigatran binds into a cavity of the proteins and forms a bidentate salt bridge to H:Asp35 and Asp189, respectively. The salt bridge is further stabilized by a conserved H-bonding pattern with L:His96 and H:Asp100D or with a water molecule and Gly219. **(C-D)** The benzimidazole moiety of dabigatran interacts by π-stacking with L:Tyr27D and Trp60D, respectively. The pyridine ring of dabigatran interacts with a T-shaped π-stacking with H:Trp52 and Trp215, respectively.

dabigatran in a hydrophobic cavity in both proteins. In the thrombin structure, this salt bridge is formed by Asp189 (chymotrypsinogen numbering) in the S1-pocket and additional H-bonds are formed by the backbone carbonyl oxygen of Gly219 and a water molecule with the amidine. In the Fab, a salt bridge is formed by H:Asp35 and a similar H-bonding pattern is formed by the ε-nitrogen of L:His96 and the backbone carbonyl oxygen of H:Asp100D. When both structures are superimposed on the benzimidazole moiety of dabigatran, additional similarities become obvious. Thrombin interacts with the benzimidazole moiety of dabigatran via parallel displaced π-stacking with Trp60D (3.5 Å distance of closest atoms) in the S2-pocket, whereas the Fab uses L:Tyr27D for the same purpose. In both crystal structures the pyridine ring of dabigatran displays a T-shaped π-interaction with a tryptophane. In the S4-pocket of thrombin this is Trp215 (3.3 Å distance), in aDabi-Fab1 it is H:Trp52. Notably, the indole moieties occupy the same space, but they are orientationally inverted due to the position of the main chains from which the Trp side chains protrude. In the thrombin structure the tryptophane protrudes from a strand that is close to the benzamidine group, whereas in the Fab from a strand that is closer to the propionic acid moiety of dabigatran. Interestingly, the N-methyl group on the benzimidazole scaffold that forms hydrophobic interactions within the S2-pocket of thrombin is not involved in binding to the Fab. Upon binding to thrombin, dabigatran buries 498 Å^2 of interface area. This is 33 Å^2 (7 Å^2) less than the interface area of the Fab after binding to the hapten. The structural similarities of aDabi-Fab to thrombin and the potent binding to dabigatran may suggest possible thrombin-like enzymatic activity of the antidote. However, a closer inspection of the binding sites of both proteins did not reveal the presence of a catalytic triad in Fab1 (see figure 14 on the following page). The binding of aDabi-Fab to thrombin substrates was furthermore analyzed by SPR (see figure 15 on the next page). Known thrombin substrates were injected over a CM5 chip with aDabi-Fab1 immobilized. None of the tested substrates, including S-2238, factors V, VIII, XIII, fibrinogen, vWf, PAR-1 or protein C bound to aDabi-Fab1 at physiological concentrations. In addition, it was demonstrated that the Fab alone did not influence clotting in the diluted thrombin time assay. Similarly, it did not shorten the clotting time in a system where prothrombin was limited and aDabi-Fab was the only protein that could potentially exert any thrombin-like activity. It also did not convert the main substrate of thrombin, fibrinogen into fibrin as measured by FPA release (see figure 15 on the following page) and it did not mimic or enhance the thrombin-mediated feedback of coagulation, resulting in elevated thrombin generation (the last two assays were performed by the department of cardiometabolic disease research. These functional results taken together therefore demonstrate that while there are certain similarities in structural elements recognizing the ligand be-

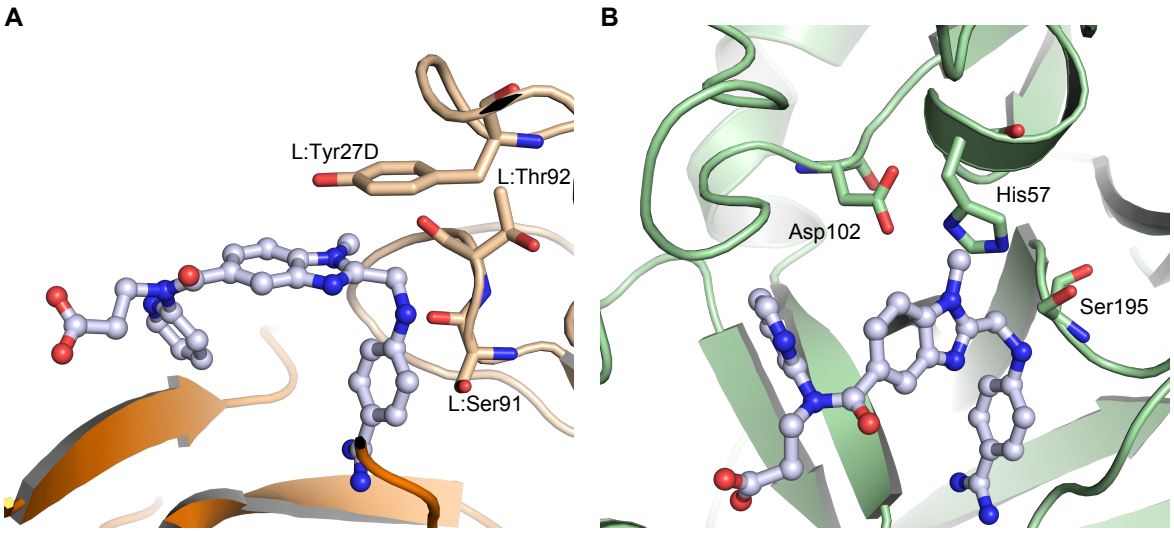


Figure 14: aDabi-Fab1 Does Not Mimic Functional Features of Thrombin. (A) The region of aDabi-Fab1 equivalent (superposition based on the benzamidine group of dabigatran) to the active site of thrombin is shown. **(B)** The catalytic triad of thrombin is formed by His57, Asp102, and Ser195. The Fab contains in the same region an arrangement of L:Tyr27D, L:Ser91 and L:Thr92 that is not able to convey proteolytic activity.

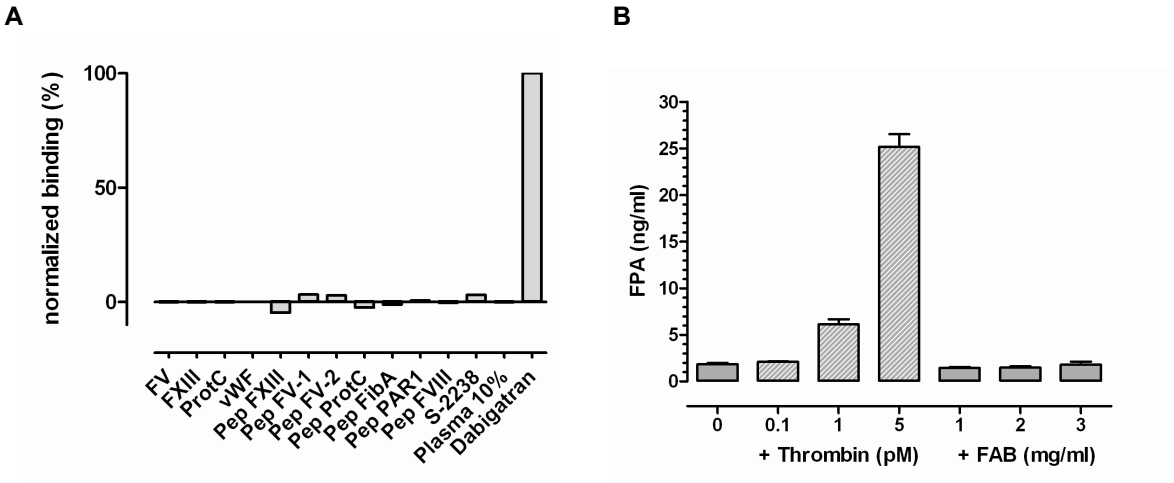


Figure 15: aDabi-Fab1 Does Not Functionally Mimic Thrombin. (A) Fab1 does not bind to substrates of thrombin in a SPR assay. **(B)** Fab1 does not cleave fibrinogen and release fibrinopeptide A (FPA), in contrast to thrombin.

tween Fab and thrombin, these do not result in any prothrombotic activity of the Fab. Additional H-bonding and hydrophobic interactions and a larger buried surface area allow for binding of dabigatran with a very high affinity of 2.1 pM, that is ~350 times stronger than the 700 pM affinity of thrombin (Wienen et al., 2007). Based on the extremely high affinity aDabi-Fab1 is able to neutralize dabigatran activity rapidly and completely. This is the first specific antidote for a next-generation anticoagulant. For patients using dabigatran, this antibody fragment will be a valuable tool to prevent bleeding in certain emergency situations.

5.2 Structure Activity Relationship Studies

The structural similarities between aDabi-Fab1 and thrombin in binding of dabigatran were unexpected. To further investigate whether there are additional similarities in the binding mode, structure activity relationship (SAR) studies were performed. Therefore a total of 13 structural homologues of dabigatran, as well as the fragment benzamidine, were tested for their binding affinity to aDabi-Fab1 and thrombin (see figure 16 on the next page). As the affinity of dabigatran for aDabi-Fab1 is extremely high (K_D = 21 pM), the KinExA method was chosen for SAR studies with the Fab. For high-affinity binders like dabigatran, a thorough affinity determination takes about 10 - 20 h. As dabigatran binds thrombin with a lower affinity (K_D = 700 pM) and as a reliable regeneration procedure was available, SPR was used for SAR studies with thrombin for convenience. Increasing compound concentrations were injected over a CM5 chip with thrombin immobilized. After each injection, the surface was successfully regenerated with 10 mM ammonium acetate, pH 4. Affinity was determined by fitting of a 1:1 binding model to kinetic data. All sensograms are shown in figure 17, K_Ds are reported in figure 21. Prior to accurate determination of the affinities of analogs of dabigatran for aDabi-Fab1, a qualitative ELISA was performed (see figure 18 on page 90). In the setting used, there is an “assay wall” for the ELISA of around 1 nM, below which no affinity can be determined. Nonetheless the ELISA allowed for a coarse binning of the analogs according to their affinities. The fragment benzamidine (BA) and compounds 12 and 1 were the worst binders in the ELISA (apparent K_D of ~1 μM), followed by analogs 13 and 2 (apparent K_D of ~10 nM). All other compounds were located close to the assay wall of 1 nM, with analogs 7, 5 and 9 displaying slightly lower affinity. Following the initial, very coarse estimation in an ELISA, the affinities were determined accurately with KinExA technology as described in section 4.8.3. The respective binding curves are shown in figure 19 and the results are summarized in figure 21 on page 94.

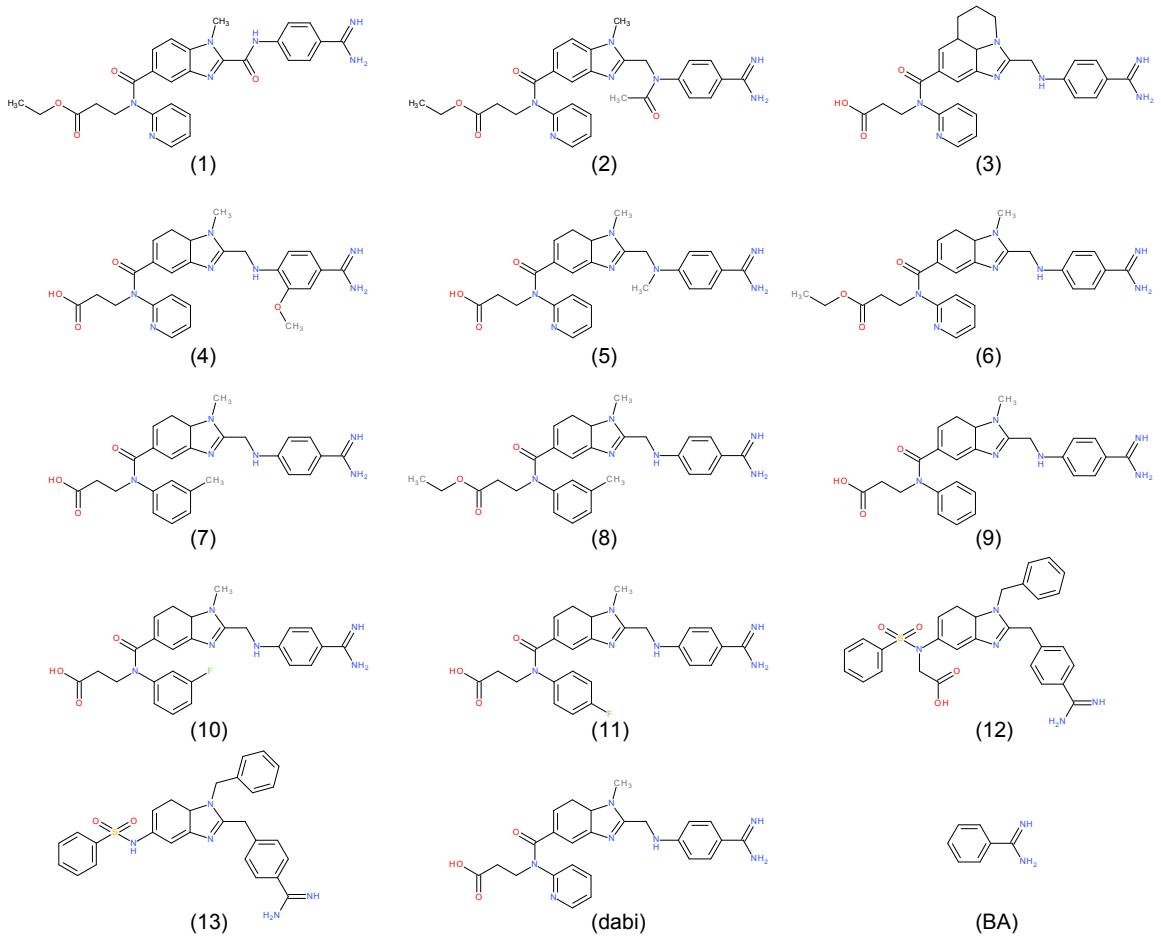


Figure 16: Chemical Structures of Dabigatran and Structural Analogs. 13 analogs were used for structure-activity-relationship studies, dabigatran (dabi) and the fragment benzamidine (BA), that were also used for SAR, are shown for the sake of completion. All analogs except 12 and 13 share the same central scaffold.

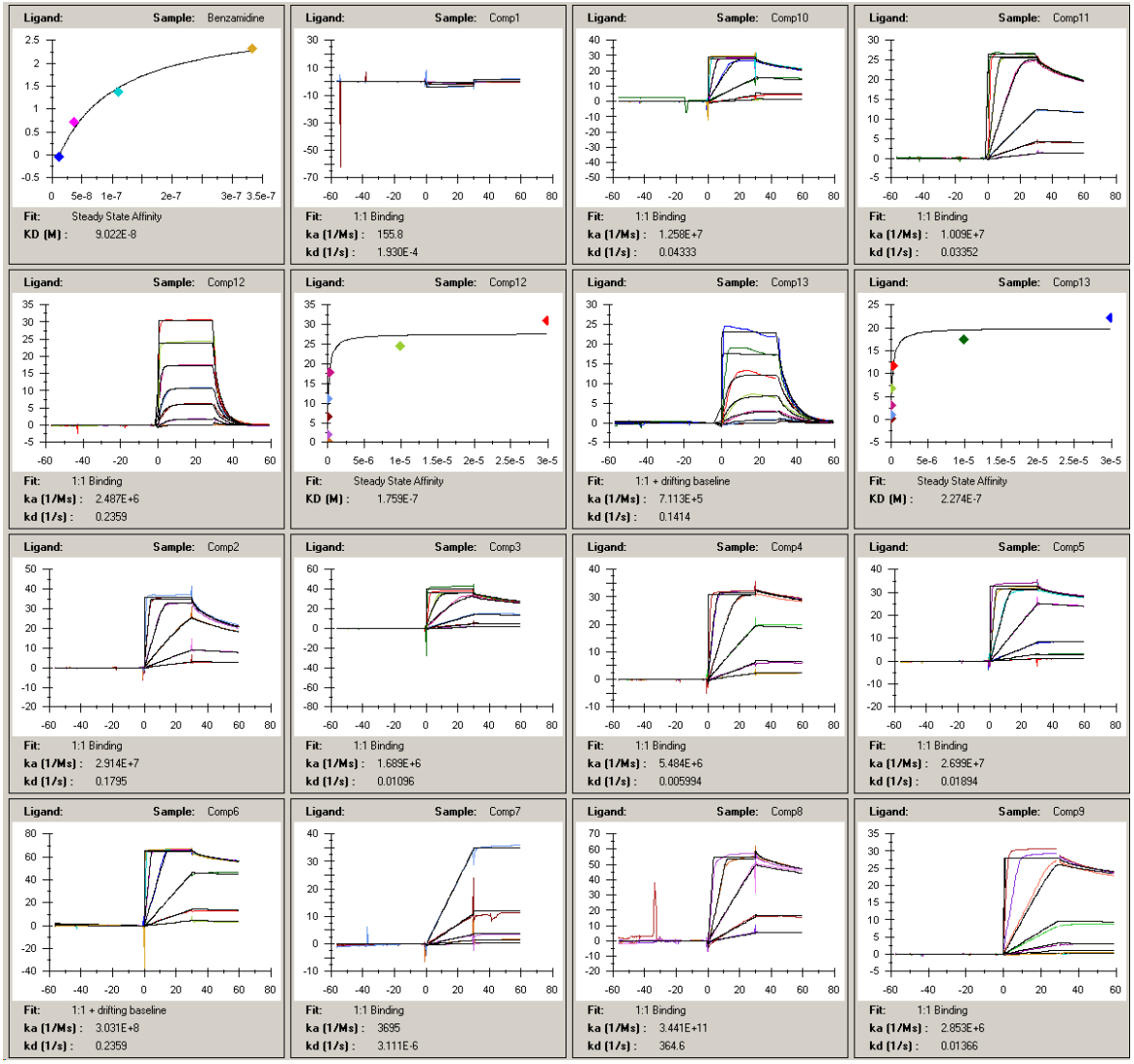


Figure 17: Affinity Determination of Dabigatran Analogs for Thrombin by SPR. Sensograms for the compounds binding to thrombin are shown. For compounds were a steady state was reached for each concentration, the equilibrium analysis is also shown. All data were fitted to a 1:1 binding model, the fit is shown as black solid line. Compound 1 showed no binding in SPR.

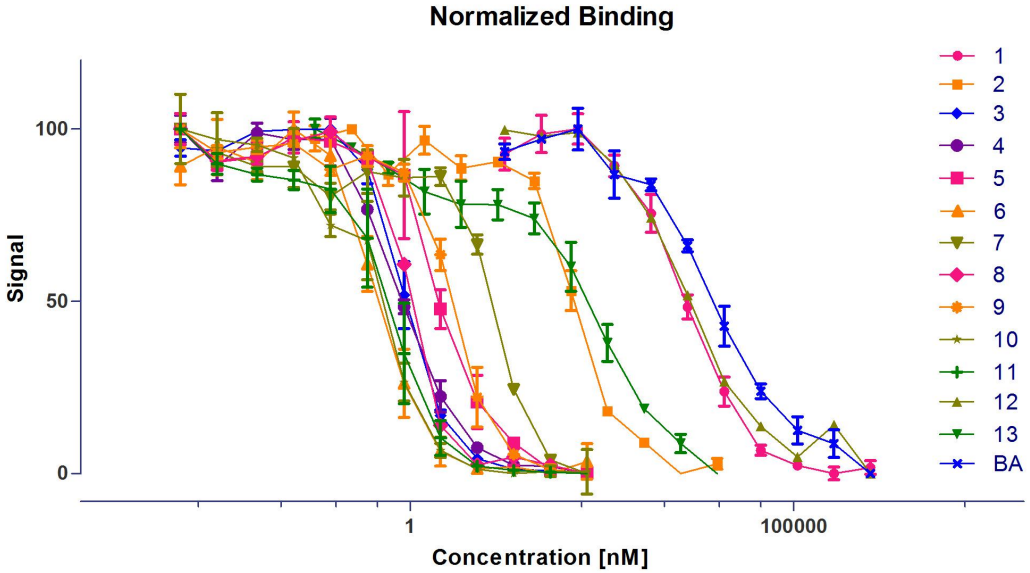


Figure 18: Affinity Binning of Dabigatran Analogs. Affinities of a panel of dabigatran analogs for aDabi-Fab1 was determined with ELISA.

5.2.1 SAR Studies for aDabi-Fab1

The crystal structures of 10 analogs of dabigatran in complex with aDabi-Fab1 were solved (see figure 20 on page 92) and the affinity of the Fab for 13 analogs was determined with KinExA technology (see figure 19 on the next page). For compound 12 no change in the fluorescence signal was observed. Compounds 12 was colored and might thus interfere with detection of the fluorescence-labeled antibody fragment. Nonetheless, Compound 12 was shown to bind to aDabi-Fab1 in the ELISA experiment (see figure 18) with a relatively low affinity, comparable to compound 1 (K_D = 34 nM in KinExA). From the panel of analogs tested, there are 5 that have a higher affinity towards aDabi-Fab1 compared to dabigatran. Compound 11 is by far the best binder with an apparent affinity of 4 fM. For such potent binders a quantitative affinity determination fails as this would exceed the instruments sensitivity. But a qualitative estimation is nevertheless possible by comparing the affinity with that of the other analogs. Compound 11 has a phenyl ring substituted with a fluoro group at para position. This fluoro group fills a hydrophobic pocket formed by L:Phe94, H:Val50, H:Thr57 and H:Gly58. The interaction of fluoro substituents with hydrophobic pockets has been shown to also increase affinity of ligands binding to thrombin (Hoffmann-Röder et al., 2006). Compounds 3, 8 and 10 have an increased interaction surface with the Fab which seems to contribute to their affinity. Compound 3 contains a tricyclic central core, whereas compounds 8 and 10 have a methyl and fluoro substituent at their phenyl rings, respectively. Interestingly the fluoro group in

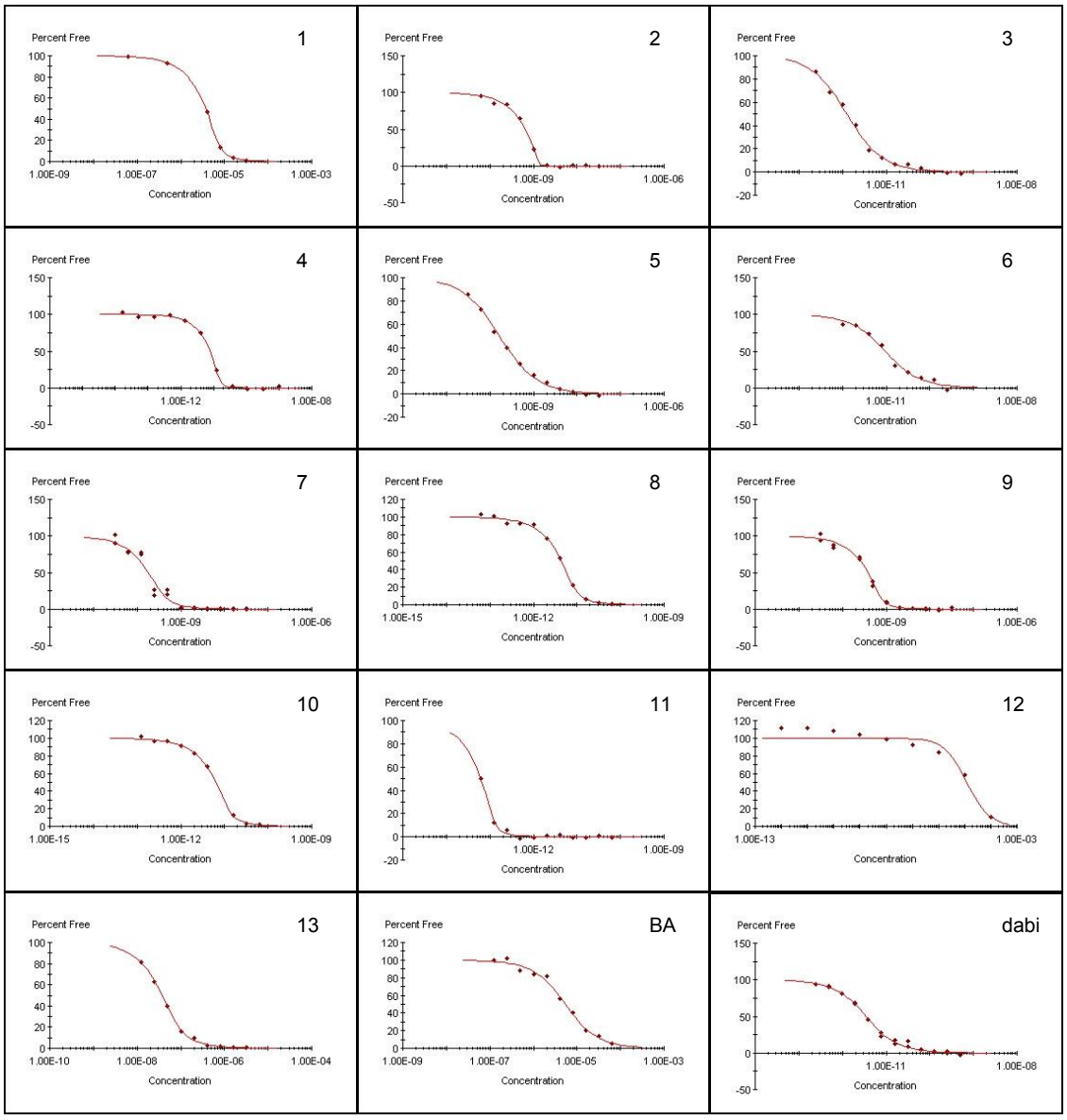


Figure 19: Affinity Data of Dabigatran Analogs for aDabi-Fab1. Affinities were determined with KinExA technology. Increasing concentrations of compound were titrated to aDabi-Fab1 of constant concentration. K_D was determined by fit of a 1:1 binding model (solid line) to the data. For detailed information see section 4.8.3.

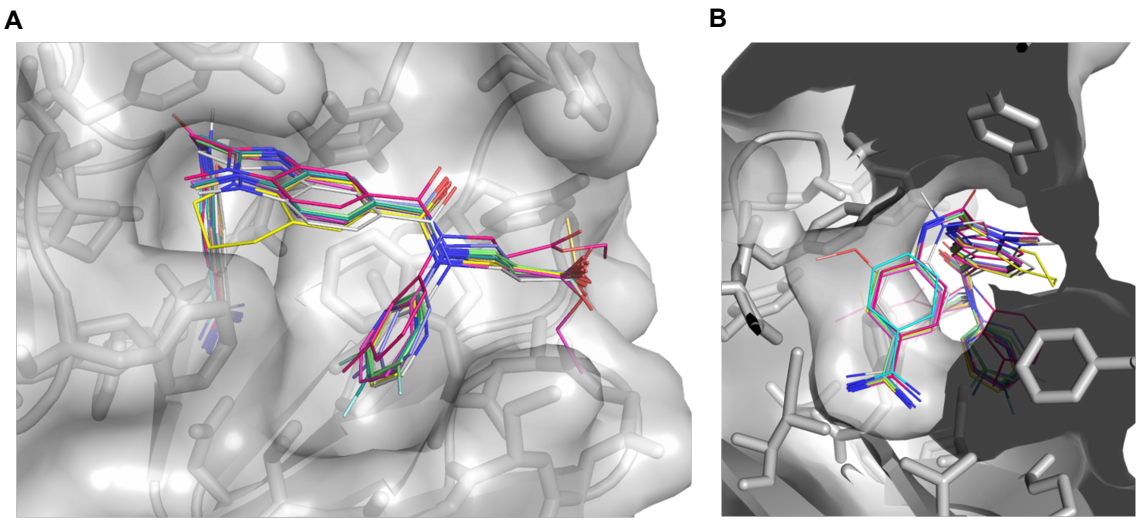


Figure 20: Crystal Structures of aDabi-Fab1 in Complex with Dabigatran Analogs. (A) Superposition of all complex structures. The increased interaction area for the tricyclic compound 3 (yellow) as well as binding of the fluoro group of compound 11 (cyan) in a hydrophobic pocket are visible. **(B)** View of the binding pocket for the benzamidine moiety. Although the methoxy group of compound 4 (red) fits well into the pocket, there is no gain in affinity for this analog compared to dabigatran.

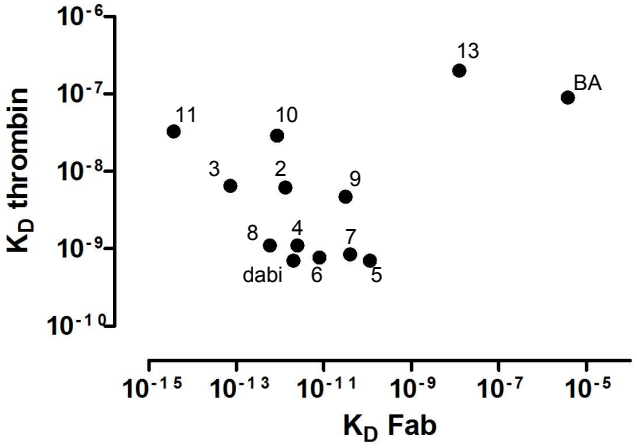
meta position increases the affinity only slightly compared to the massive effect of the same substituent in para position observed for compound 11. An alternative conformation for the fluorophenyl was observed where the fluoro group pointed away from the Fab and was solvent exposed. As both conformations were equally populated this finding further indicates that a fluor atom in meta position has no major effect on the affinity to aDabi-Fab1. Compound 2 binds with an affinity close to that of dabigatran, although it has a acetyl substituent at the exocyclic nitrogen that would clash with L:Tyr32. Binding of compound 2 might induce conformational rearrangements which might be compensated by favorable interactions with the compound, but this remains hypothetical as it was not possible to crystallize this complex. Compound 4 binds the Fab with virtually the same affinity as dabigatran. This is to some extent unexpected as the methoxy group located at the benzamidine ring tightly fits into a hydrophobic cavity of the Fab as seen in the crystal structure. But the loss in entropy of this group upon binding might antagonize a possible gain of affinity that comes with larger interaction area. Compound 6 is the ethyl ester of dabigatran and binds with a slightly lower affinity of 8.0 pM to the Fab. As the propionic acid moiety does not form any interactions with the Fab, this difference may be due to impurities of this compound. It remains elusive why compounds 7 and 9 have binding affinities that are one order of magnitude lower than dabigatran, especially as compound 7 is a hydrolyzed variant of the ester of compound 8. The crystal structure of compound 5 in complex with the Fab offers an explanation for its reduced affinity of 110 pM.

The additional methyl group compared to dabigatran forces the side chain of L:Tyr32 to move away from the binding site. In addition the NH-π interaction formed by dabigatran is no longer present. Furthermore the whole compound is positioned slightly displaced compared to all other costructure which might indicate an energetically less favored conformation. Compounds 1, 12 and benzamidine have the weakest affinity to aDabi-Fab1. Benzamidine is a small compound fragment of 120 Da and fragments are known to be weak binders (Erlanson et al., 2004). Compound 13 (as well as 12) has a core different to the other compounds. These compounds have for example a methylene linker between benzamidine and benzimidazole. Although no cocrystals were available, one can imagine that both compounds will not fit to the binding cavity without major conformational rearrangements of the Fab. Although a cocrystal structure is available for compound 1, the low affinity of 340 nM for the Fab is less clear. The carbonyl oxygen in the linker is in close contact (3.3 Å) with three carbon atoms of L:Tyr27D, L:Tyr32 and L:Thr92. The location of the negative partial charge of the oxygen in an otherwise hydrophobic pocket might be energetically unfavored but it remains doubtable if this suffices for a affinity reduction of 5 orders of magnitude compared to dabigatran. Nonetheless, compound 1 was also one of the weakest binders in the preceding ELISA.

5.2.2 Correlation of Affinity for aDabi-Fab1 and Thrombin.

The structural similarities between thrombin and aDabi-Fab1 regarding binding of dabigatran are striking. To investigate possible similarities in binding to structural analogs of dabigatran a correlation plot was drawn. As visible in figure 21, there is no obvious correlation in binding to the two proteins. Only for Compound 13 and for benzamidine there is some correlation as these are the weakest binders. But as benzamidine is a small fragment of dabigatran and compound 13 has a central framework that differs significantly from all other analogs this finding is unsurprising. Besides the features explained in section 5.1.8, aDabi-Fab1 and thrombin share no obvious structural similarities. The proteins on the whole have a different binding site topology. The binding pockets found in thrombin are different from those found in aDabi-Fab1 and except the bidentate salt bridge in a hydrophobic pocket, there is no common pattern in electrostatics or H-bond donors and acceptors. The lack of correlation in binding was thus to some extent expected and verified by the SARs.

A



B

Compound	K_D thrombin [M]	K_D aDabi-Fab1 [M]
1	n.d.	3.4E-07
2	6.2E-09	1.3E-12
3	6.5E-09	7.0E-15
4	1.1E-09	2.5E-12
5	7.0E-10	1.1E-10
6	7.7E-10	8.0E-12
7	8.5E-10	4.0E-11
8	1.1E-09	5.9E-13
9	4.9E-09	3.1E-11
10	2.9E-08	8.5E-13
11	3.3E-08	4.0E-15
12	9.5E-08	n.d.
13	2.0E-07	1.3E-08
BA	9.0E-08	3.8E-06

Figure 21: Structure Activity Relationship of Dabigatran Analogs. (A) Correlation plot of affinity of analogs to aDabi-Fab1 and thrombin. **(B)** Explicit values for the affinity of the analogs is given in the table. Affinities for thrombin were determined by SPR, for aDabi-Fab1 by KinExA.

5.3 aDabi-Fab2 and 3

5.3.1 Purification of aDabi-Fab2/3

aDabi-Fab2 and 3 were expressed in HEK cells and purified by affinity chromatography on benzamidine sepharose as described in section 4.3.6. A representative SDS-PAGE and chromatogram of the purification is shown in figure 22. Both Fabs were used for ensuing biophysical studies and X-ray crystallography.

5.3.2 Crystallization of aDabi-Fab2/3

The complex of aDabi-Fab2/3 with dabigatran was formed by incubation of the protein with 2.5 molar excess of dabigatran for 1 h at 4 °C. Crystallization trials were set up using the vapor diffusion method with the sitting drop procedure. 200 nl of the protein was mixed with an equal volume of screening solution and equilibrated against the screening solution. aDabi-Fab2 in complex with dabigatran crystallized in 100 mM SPG buffer (pH 6) and 25% PEG 1500. Crystals appeared after 8 days of incubation at 20 °C. aDabi-Fab3 in complex with dabigatran crystallized in 100 mM SPG buffer (pH 7) and 25% PEG 1500. Crystals appeared after 8 days of incubation at 20 °C. Both uncomplexed Fabs crystallized at 100 mM PCB buffer (pH 6) and 25% PEG 1500. Crystals appeared after 1 day of incubation at 20 °C. Crystals were cryoprotected by soaking them in the crystallization solution supplemented with 15% glycerol and flash frozen in liquid nitrogen at 100 K.

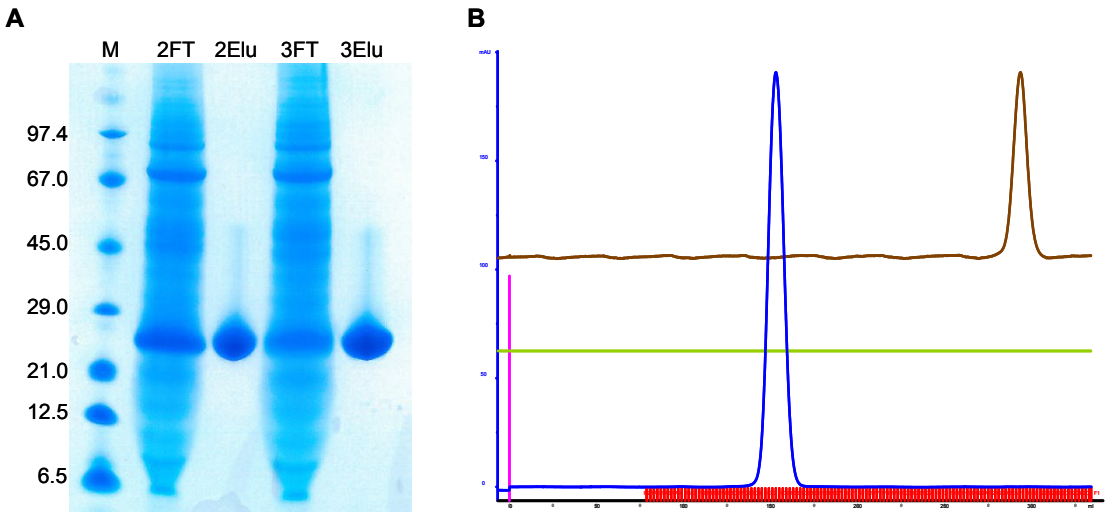


Figure 22: Purification of aDabiFab2 and 3. (A) SDS gel of Fab affinity purification with benzamidine sepharose (2/3FT: flow-through of Fab2 and 3; 2/3Elu: Elution of Fab2 and 3). **(B)** Chromatogram of final polishing of Fab2 on SEC. The chromatogram for Fab3 is virtually identical.

5.3.3 Overall Structure of the Fabs

The crystal structure of the antigen binding fragments aDabi-Fab2 and 3 that bind the drug dabigatran were solved in their bound and unbound forms (for detailed crystallographic information of all structures see appendix A) at resolutions of 2.2 Å or better. Both Fabs differ only at 2 positions in the framework regions 3 and 4 of the light chain and show no major structural differences besides these two amino acids. For convenience the further structural description will refer to Fab2. If the respective features differ significantly between two protomers in a unit cell or between the different crystal systems, they will be mentioned explicitly. For all structures the electron density for the major part of the model is very well defined. For the apo structure of aDabi-Fab2 a stretch in framework region 2 of the heavy chain from H:Ala40 to H:Gly44 (numbering according to Wu and Kabat (Johnson and Te Wu, 2001)) is weakly defined and shows different backbone conformations between the two protomers in the asymmetric unit of the crystal, which indicates an intrinsic flexibility for this loop. In the apo structure of aDabi-Fab3 the constant domain of the heavy chain is not defined for the amino acids H:Ser127 to H:Gly133. In the complex structure of aDabi-Fab2 that crystallized in space group $P2_1$ the intermolecular disulfide bridge adopts a different conformation between the two protomers. The Fab is composed of two variable domains ($V_L\varkappa$ and V_H) and two constant domains ($C_L\varkappa$ and C_HIgG1). These domains show the typical immunoglobulin fold (Bork et al., 1994). The antigen binding site of antibodies is formed by the complementarity-determining regions (CDRs) on the surface of the variable domains which are named from L1 to

H3. The angle between the two pseudo-dyad axes relating the two variable and constant domains is called the elbow angle. All Fabs presented here display an elbow angle of 127.7 – 129.2°, only the apo form of aDabi-Fab3 has a larger elbow angle of 157.8°. This difference is probably due to crystal packing effects, as noticed previously (Stanfield et al., 1993). Noteworthy the benzamidine moiety of dabigatran is bound identical in all complex structures with a rmsd of 0.031 Å over all atoms, whereas the rest of the hapten adopts slightly different conformations with a maximum rmsd of 1.576 Å. Usually the CDR loops adopt conformations that can be grouped into canonical structural classes (Al-Lazikani et al., 1997). CDR:L1 ranges from residue 24 – 34 of the light chain and has a 5-residue insertion at position 27. It belongs to canonical structure class 4. The residues 50 – 56 build CDR:L2 that resembles class 1. Like in all antibody structures published, the amino acid on the tip of this loop (L:Val51) adopts a strained conformation and contributes to the formation of a γ-turn. The third CDR in the light chain (residues 89 – 97) has a conformation typical for canonical structure 1. CDR:H1 spans residues 26 – 35 and is similar to the canonical structure class 1. The CDR:H2 loop adopts canonical class 1 with a proline inserted at position 52. For CDR:H3 that displays an insertion of 2 amino acids at position 100, no classical canonical structure can be assigned. The region of CDR:H3 proximal to the framework is called the torso. Depending on the conformation, the torso can be classified as bulged or non-bulged (Morea et al., 1998). All Fabs described here adopt a bulged conformation the characteristic structural feature of which is the H-bond between the sidechain nitrogen of H:Trp103 and the backbone carbonyl oxygen of the residue 3 positions N-terminal to it, in this case H:Phe100B. The β-bulge is stabilized by H-bonds formed between the sidechain of H:Asp101 and the backbone nitrogen and the hydroxyl group of H:Thr96, respectively. The light chain further contributes to stabilize the bulge. L:Tyr36 forms an H-bond to the backbone carbonyl oxygen of H:Phe100B in the apo structure and to its backbone nitrogen in the complex structure, respectively.

5.3.4 Binding Mode of the Fabs

The dabigatran binding site of the Fab is located at the interface of the variable domains (see figure 23). All CDR loops except L2 are involved in dabigatran binding. The binding is mediated by hydrophobic interactions, H-bonds and a salt bridge. The benzamidine moiety of dabigatran inserts into a cavity formed by heavy and light chain, whereas the benzimidazole, carboxamide and pyridine moieties are partially exposed to solvent. The fact that the carboxylate moiety is not involved in binding to the Fab is consistent with the employment of the hapten conjugated at this position

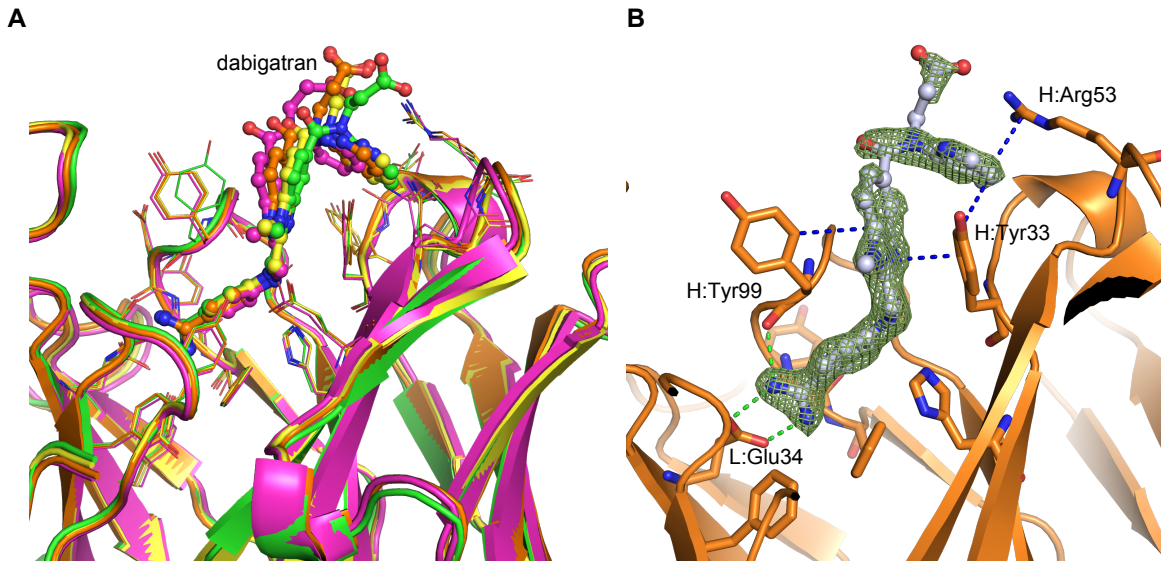


Figure 23: Crystal Structures of aDabi-Fab2/3 in Complex with Dabigatran. (A) Superposition of 4 crystallographic independent complex structures (dabigatran colored as the respective protein). Except for the benzamidine moiety dabigatran adopts slightly different conformations. H:Tyr and H:Tyr also adopt different sidechain conformations but. **(B)** Detailed view of the interactions of aDabi-Fab2 with dabigatran. Electron density of dabigatran is shown at a contour level of 1.5 σ. H-bonds are shown as green dotted lines, π-stacking interactions in blue.

to a carrier protein for immunization. As dabigatran adopts slightly different conformations in the complex, the contact area (defined as the difference in total accessible surface area divided by two) between dabigatran and the Fabs also varies between 468 Å and 492 Å. This variation is caused by the contacts formed to the heavy chain of the Fab as it differs between 370 Å and 392 Å, whereas the contacts for the light chain remain constant at about 100 Å.

Hydrophilic Interactions The most important interaction for binding of dabigatran to the Fabs is a bidentate salt bridge formed at the interface region of heavy and light chain. The benzamidine moiety of dabigatran inserts into a cavity and there interacts with L:Glu34 in CDR:L1 with a nitrogen-oxygen distance of 2.7 Å. This interaction is additionally stabilized as one of the nitrogens forms a H-bond to the carbonyl oxygen of H:Tyr99 (2.6 Å) in CDR:H3. These polar interactions are expected to contribute strongly to the free energy of binding as they reside in an otherwise hydrophobic environment provided by the side chains of H:Phe100B, L:Tyr32, L:Tyr89 and L:Ala91, which provide efficient solvent shielding. The amine nitrogen and the aldimine nitrogen of dabigatran both participate in a H-bonding network which is mediated by two water molecules. These waters bind to H:Tyr33 and H:Gly95, but as they are solvent exposed these H-bonds are not thought to contribute significantly to binding affinity. In three of the four protomers of the complex structures solved,

the carboxyl group of dabigatran forms an H-bond with H:Arg53.

Aromatic Interactions Parallel displaced as welt as T-shaped π-interactions are observed in the costructures of dabigatran binding to its neutralizing Fabs. In CDR:H3 H:Tyr99 forms a edge-to-face interaction with the benzimidazole moiety of dabigatran. Due to the different conformations of dabigatran and the fact that H:Tyr99 is involved in a crystal contact in 3 protomers, the distance of the closest atoms of the ring systems varies between 3.8 and 4.4 Å. H:Tyr33 is involved in displaced π-stacking with the benzimidazole moiety (3.3 – 3.4 Å) and interacts edge-to-face with the pyridine ring of the hapten (3.5 – 3.8 Å). This pyridine group forms an additional cation π interaction to the guanidinium group of H:Arg53 in CDR:H2 with a distance of 3.8 – 4.4 Å. Arginine is the predominant amino acid that provides the cation in cation π interactions (Gallivan and Dougherty, 1999). In the binding pocket for the benzamidine moiety, H:Phe100B displays a N-H/π interaction with an amidine nitrogen (3.4 Å distance of nitrogen to ring centroid). In this nonclassical H-bond the π-electrons of the phenyl ring act as H-bond acceptor.

5.3.5 Apo Structures of the Fabs

The apo structures of the two Fabs aDabi-Fab2 and aDabi-Fab3 were solved at a resolution of 1.8 and 2.1 Å, respectively. Since aDabi-Fab2 contains two Fabs per asymmetric unit the two structures provide three independent views of the ligand-free binding site. In all structures the electron density is well defined for the major part of the model, especially for the paratope region. The CDRs adopt the same canonical classes as observed in the structures of the dabigatran complexes. But the distal part of CDR:H3 adopts a different conformation in both Fabs compared to their complexed forms. Its backbone orientation changes, which results in blocking of the dabigatran binding site, especially L:Glu34 is no longer accessible in the apo structure. The stretch from H:Ser97 – H:Tyr100A changes its backbone conformation which is most pronounced for H:Asp100 that has a maximum distance of 5.2 Å between the apo and complexed structure, concerning its Cα atoms and 10.2 Å for the carboxyl oxygen. The change in side chain orientations of H:Tyr99, H:Asp100 and H:Tyr100A results in a blocked binding site. The side chain of H:Tyr100A completely blocks L:Glu34, of which the carboxyl group is rotated by 53°. In addition the side chain and the carbonyl group of H:Asp100 are located in the cavity that is occupied by the benzamidine moiety of dabigatran in the complex structure.

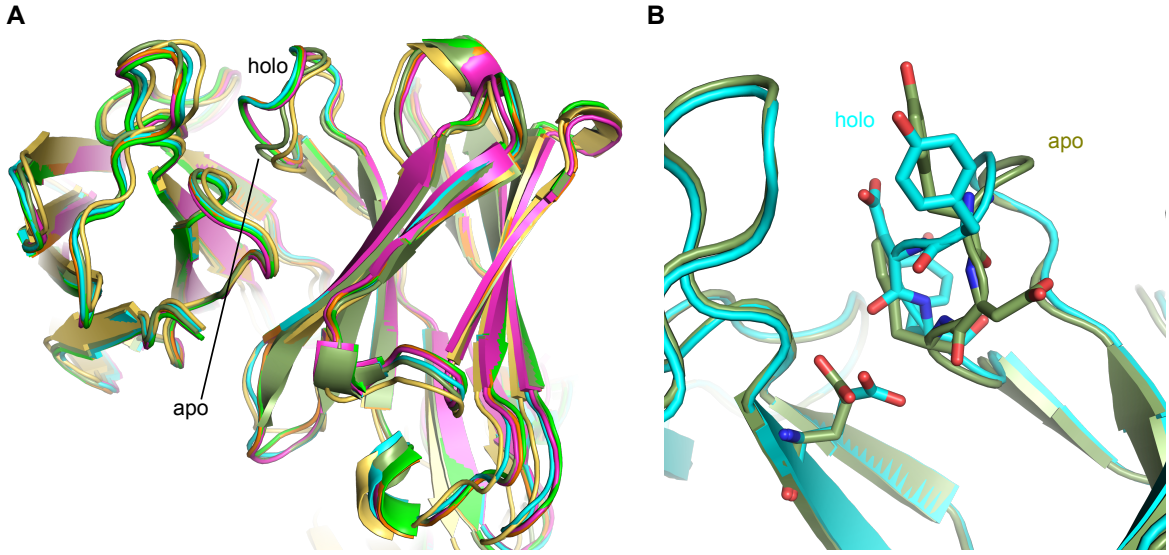


Figure 24: Conformational Changes Between Apo and Complex Structures of aDabi-Fab2/3. (A) Superposition of apo and holo conformations of the binding site. Only CDR:H3 adopts a different conformation in presence of the ligand. **(B)** Detailed view of CDR:H3. H:Asp100 and H:Tyr100A block access to the hot spot L:Glu34 in the apo structures.

5.3.6 Reduced Affinity Compared to aDabi-Fab1

aDabi-Fab1 mimics structural features of thrombin (see section 5.1.8) regarding the binding of dabigatran. aDabi-Fab2/3 mimics the interactions with dabigatran made by thrombin to a much lesser extent. In the thrombin-dabigatran complex the benzamidine group forms a bidentate salt bridge with Asp189 in the S1-pocket and additional H-bonds to the carbonyl oxygen of Gly219 and a water molecule. In aDabi-Fab2/3 a salt bridge is formed to L:Glu34 with an additional H-bond to the carbonyl oxygen of H:Tyr99, but in contrast to the Fab1 and thrombin complex there is no further H-bond in the binding cavity. In the thrombin costructure, the benzimidazole moiety of dabigatran is bound by parallel displaced π-stacking with Trp60D in the S2-pocket. aDabi-Fab1 as well as aDabi-Fab2/aDabi-Fab3 use a tyrosine for the same purpose. The pyridine ring of dabigatran forms an T-shaped π-interaction with Trp215 in the S4-pocket of thrombin and with H:Trp52 in aDabi-Fab1, respectively. In both cases the partially positively charged hydrogens at positions 3 and 4 of the pyridine ring interact with the π-electrons of tryptophane. There is no such interaction in the costructure with aDabi-Fab2/aDabi-Fab3. Here the pyridine moiety forms a cation-π interaction with the guanidinium group of H:Arg53 and an edge-to-face π interaction with H:Tyr33. In the complex structure H:Tyr99 displays a poor conformation according to MolProbity (Chen et al., 2009) whereas in the apo form it adopts a rotamer seen in 82% of all reference structures. In the complex with dabigatran L:Glu34, that resides at the bottom of the binding cavity, adopts a side chain conformation which

occurs at a frequency of only 1% in the reference structures. In the apo structures L:Glu34 displays a side chain conformation that is more frequently observed in 12% of the known structures and therefore thought to be energetically favored.Compared to aDabi-Fab1, Fab2 and 3 form fewer H-bonds, a less tight network of hydrophobic interactions and some of the amino acid side chains adopt conformations that are energetically unfavored. This offers a qualitative explanation for the lowered affinity of the Fabs compared to Fab1.

5.3.7 Conformational Changes Between Apo and Complex Structure

In the complexed structures of aDabi-Fab2/3 the benzamidine moiety of dabigatran resides in a pocket that is formed by the hydrophobic side chains of H:Tyr99, H:Phe100B, L:Tyr32, L:Phe89, L:Ala91 and L:Tyr96. This hydrophobic environment allows for binding to the benzamidine ring and energetically favors the formation of a salt bridge between the amidine group and L:Glu34 at the bottom of the pocket. In the apo structure this pocket is blocked by CDR:H3 that undergoes major conformational rearrangements compared to the complex structure. For antibody fragments where the unbound and complexed structures have been solved, a variety of conformational changes has been reported ranging from side chain movements (Niemi et al., 2010) and small shifts of CDRs (Leysath et al., 2009; Stanfield et al., 2006) and domain motions (Pozharski et al., 2004) to more pronounced effects in the backbone conformations of CDRs as reported by Rini (Rini et al., 1992) and James (James et al., 2003) for example. Although CDR:H3 adopts different conformations in the apo and complexed structures presented here, this does not necessarily suggest that CDR:H3 is highly flexible. As the same backbone conformations are present in different crystallographic environments one can expect these conformations to be an energetic minimum in the presence and absence of dabigatran, respectively. This assumption is supported by MD simulations: Upon removal of dabigatran from a complex structure of aDabi-Fab2, CDR:H3 adopted conformations comparable to those seen in the apo crystals of Fab2 and 3. Although the apo conformation was not completely adopted, it was approached during MD (see figure 25). A plot of the first two eigenvectors that contain majority of information about structural rearrangements, clearly shows the approximation towards the apo conformation. In the absence of the hapten CDR:H3 seems to blocks to a binding hotspot consisting of a charged amino acid (L:Glu34) in an otherwise hydrophobic environment. It is likely that the formation of a salt bridge between L:Glu34 and the benzamidine moiety of dabigatran is the major binding force. As the sidechain of L:Glu34 is inaccessible in all the apo structures, it is unlikely that from an initial encounter complex dabigatran could have induced a con-

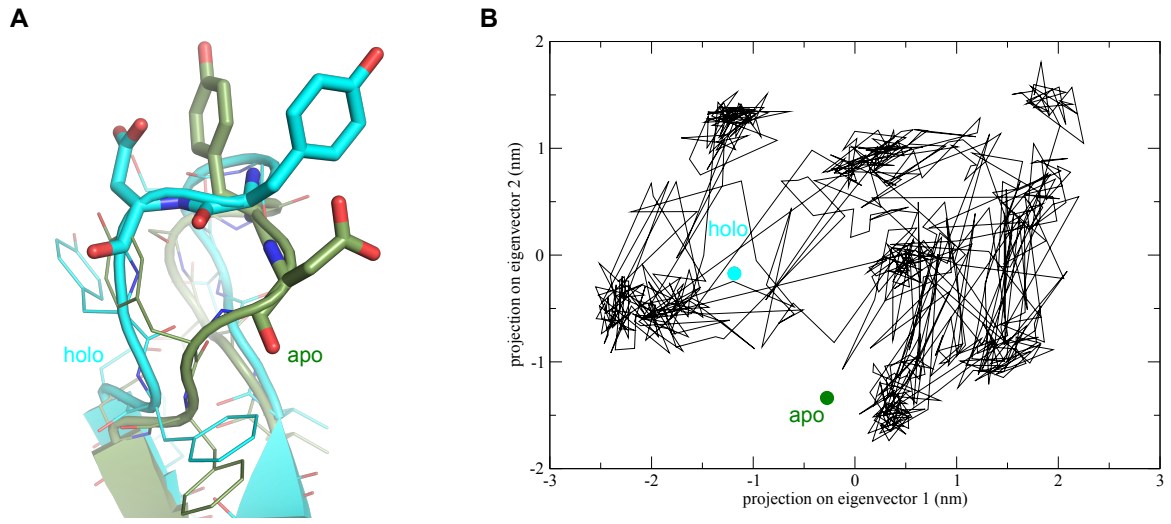


Figure 25: Molecular Dynamics Simulation of CDR:H3 After Removal of Dabigatran From the Holo Structure. (A) Conformations of CDR:H3 as seen in the apo and holo X-ray structures of aDabi-Fab2. **(B)** Projection of the two major eigenvectors of movement of CDR:H3 during a MD simulation of 500 ns. Starting from an energy-relaxed holo structure an approximation towards the apo structure is visible. This further strengthens the hypothesis that that the conformation seen in the various apo structures is present in solution and not caused by crystal contacts.

formation as observed in the complex structure. a MD study suggested, that in case of antibody hapten binding, the “conformational selection” pathway, where binding of the hapten leads to a redistribution of already pre-existing conformational substates, is favored over the “induced-fit” mechanism (Okazaki and Takada, 2008). MD simulations were performed for the apo form of Fab2. Here the holo conformation was not approached, but CDR:H3 rather opened access to the binding hotspot. This was expected as the conformation of CDR:H3 seen in the complex structure is stabilized by interactions with dabigatran (H-bond between amidine and the carbonyl oxygen of H:Tyr99, edge-to-face interaction of benzamidine moiety with the sidechain of H:Tyr99) So instead of an “induced-fit” mechanism (Koshland Jr, 1958), in this case it seems more reasonable to assume a “conformational selection” followed by “induced fit” model of binding (Kumar et al., 2000). Both mechanisms have previously been reported for antibody antigen complexes (James et al., 2003; Estabrook and Reich, 2006; Rini et al., 1992; Wilson and Stanfield, 1994).

5.3.8 Thermodynamics of Binding of Dabigatran to Fabs and Thrombin.

Isothermal titration calorimetry was performed as described in section 4.8.2 to determine the enthalpic and entropic contributions in binding of dabigatran to thrombin and aDabi-Fab1 and 2. All data was fit to a 1:1 binding model. From a isothermal titration calorimetry experiment the binding affinity can be assessed from the slope

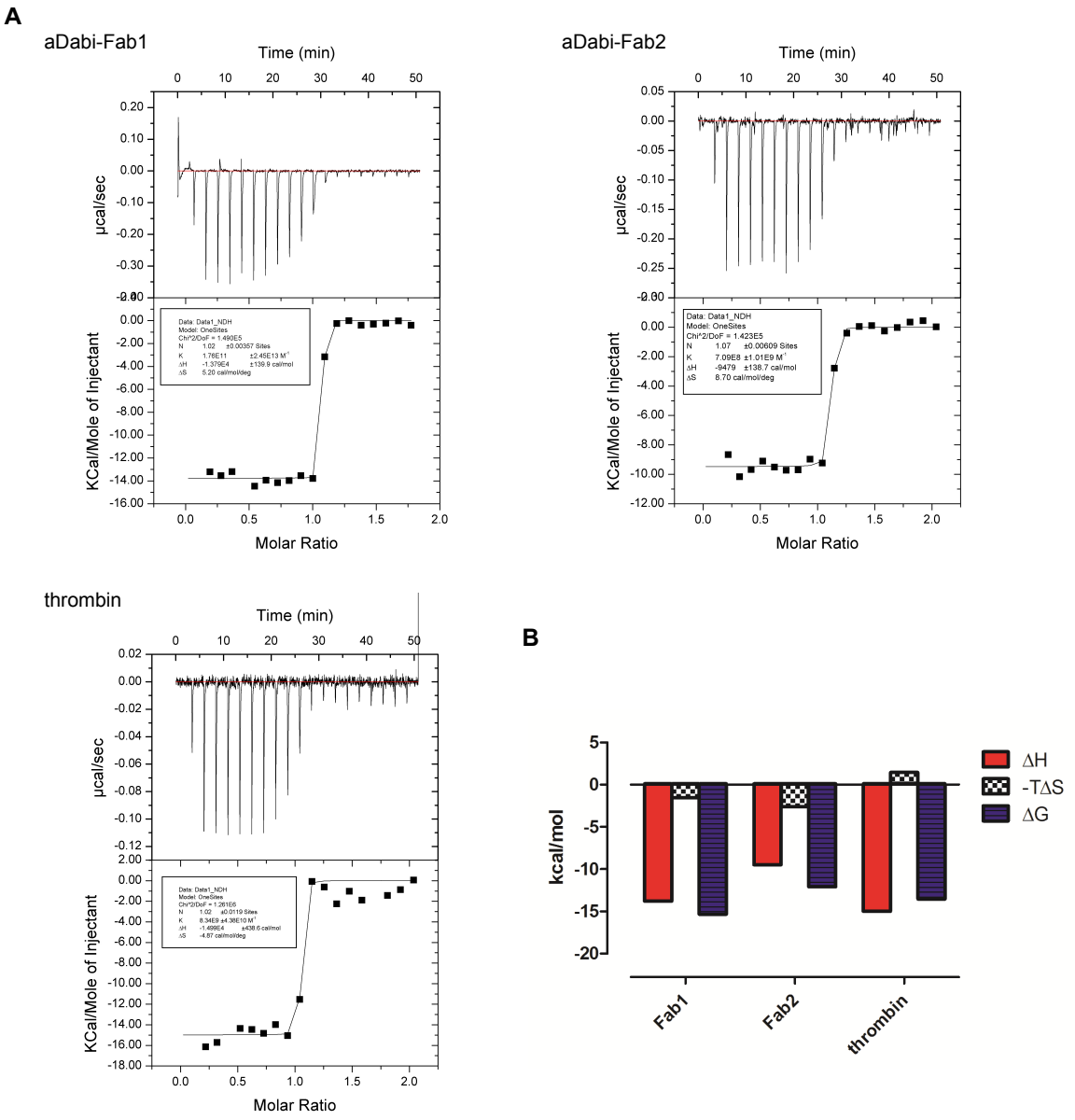


Figure 26: Thermodynamics of Dabigatran Binding. (A) Raw power signal and corresponding thermogram for binding of dabigatran to aDabiFab1, Fab2 and thrombin. The slope of the thermogram is too steep to quantitatively determine K_D, but a qualitative statement on the contribution of entropy and enthalpy can be given. **(B)** Thermodynamic signatures of dabigatran binding to the three proteins. Enthalpy is the dominating term in all cases.

of the thermogram. As visible from figure 26 on the preceding page, the slope of all sensograms is too steep to be evaluated reliable, thus only a quantitative analysis is given here. Modern ITC devices are capable of determining affinities of up to the one-digit nanomolar scale. With displacement calorimetry where a high-affinity ligand is titrated into protein that is prebound to a weaker inhibitor it would be possible to extend the sensitivity of an ITC experiments (Sigurskjold, 2000). But this method relies on the availability of binders that significantly differ in their thermodynamic binding signature, which were not available in the case of dabigatran. Nonetheless the binding affinities for thrombin (120 pM determined by ITC), aDabiFab1 (K_D = 5.7 pM) and aDabi-Fab2 (K_D = 1.4 nM) for the proteins match quite well with the values determined by other techniques that are more accurate for high affinity binders. The entropy term is small for all three binding events and favors binding to the Fabs but not to thrombin. More interestingly the binding is in all cases dominated by the enthalpy term, indicating that H-bonds and van der Waals interactions contribute the majority of binding affinity. Compared to aDabi-Fab2, aDabi-Fab1 has larger enthalpic contribution to binding (13.8 kcal/mol vs 9.5 kcal/mol) which can be explained by the 2 additional H-bonds seen in the crystal structures. Regarding thermodynamics dabigatran binds to thrombin and the Fabs in a comparable way, with a dominating enthalpy term.

6 Results and Discussion - Antibodies Against PCSK9

PCSK9 is a protein involved in the pathological state of dyslipidemia. It binds to the LDL receptor (LDL-R) and thereby prevents its recycling. Reduced levels of LDL-R result in increased plasma concentrations of LDL cholesterol (LDL-C). LDL-C can accumulate in blood vessels leading to atherosclerosis which associated with severe implications like peripheral artery occlusive disease, coronary heart disease and stroke. In this section three antibodies against PCSK9 are presented that were identified to block internalization of LDL-R in a cellular assay. mAb1 binds a linear epitope near the EGF-A binding site of PCSK9. Its costructure with a peptide representing PCSK9 was solved and the individual contributions of each amino acid to binding affinity were determined with SPR analysis. mAb2 binds to the C-terminal domain of PCSK9 and does not inhibit interaction with LDL-R. mAb is the first antibody binding to the CTD of PCSK9 that lowers LDL-C levels *in vivo* in nonhuman primates. For mAb3 the epitope could not be determined as this antibody was difficult to crystallize. Only a truncated version of Fab3, that was missing CDR:L2 crystallized and formed intermolecular contacts that are not possible for intact Fabs.

6.1 Purification of PCSK9

Full-length PCSK9 and PCSK9ΔCTD lacking the C-terminal domain were expressed in insect cells. Both constructs were purified by IMAC and SEC as described in section 4.3.1 and 4.3.2. For both protein variants an exemplary SDS gel and a chromatogram of the final polishing step with SEC is shown in figure 27. Both proteins were of high purity after SEC. In the SDS gels the upper band of 57 kDa represents the catalytic and C-terminal domain. The lower band represents the prodomain, that obviously contains a second site of cleavage as visible from the faint additional band with a slightly lower molecular weight. The truncated species could be separated from the native protein by anion exchange chromatography, but during this step about 50% of the protein were reproducible lost. As both species crystallized in the same conditions and were not distinguishable in binding assays, anion exchange was not further investigated in this work. Both proteins were subsequently used for binding studies or for complex formation with various antibody fragments.

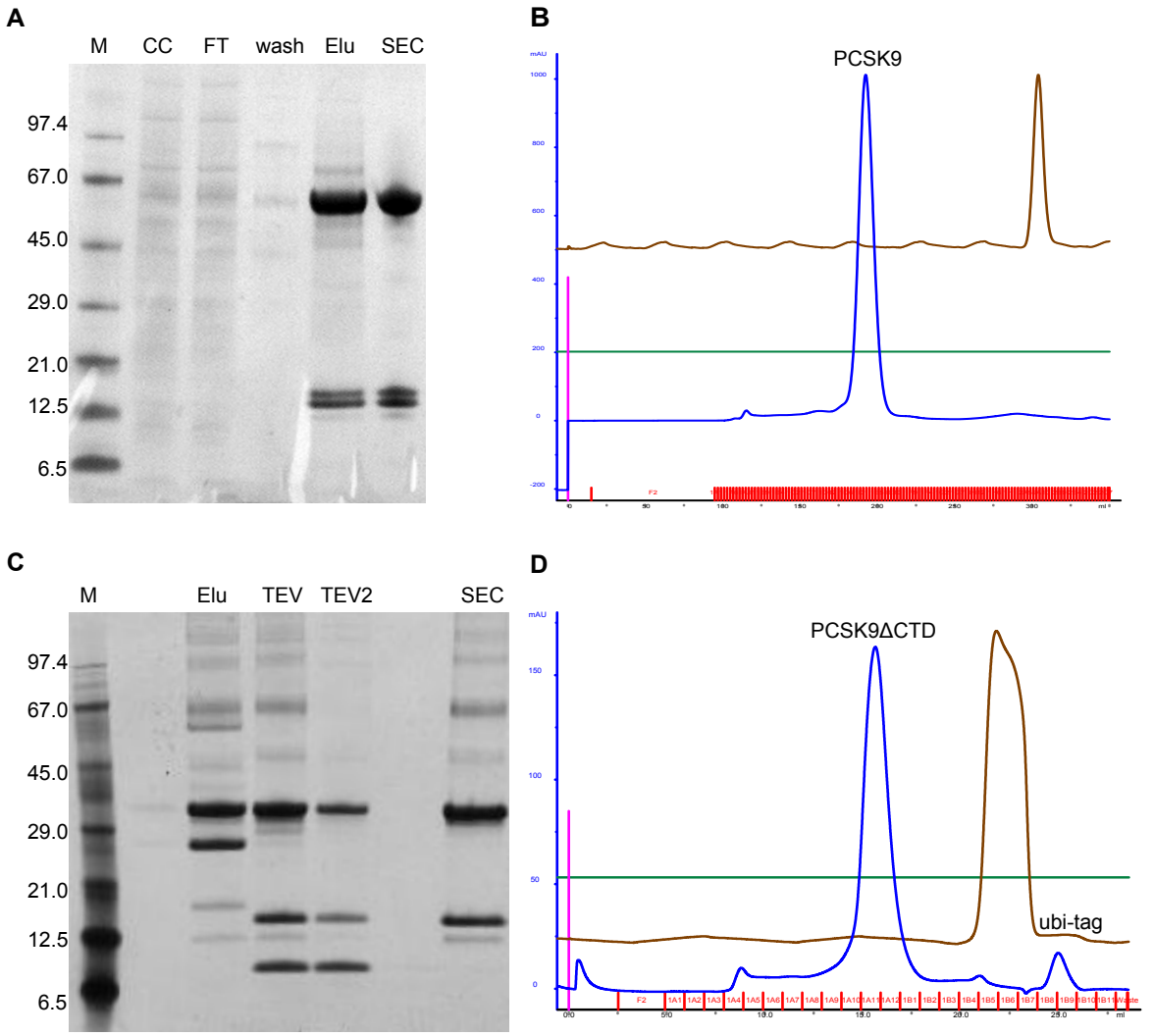


Figure 27: Purification of PCSK9 Proteins. PCSK9 purification is shown in the upper, PCSK9ΔCTD purification in the lower panel. **(A)** SDS-PAGE of PCSK9 purification (M: marker; CC: cell culture medium; FT: flow-through of IMAC; wash: wash fraction of IMAC; Elu: elution fraction of IMAC; sec: fraction from major protein peak after SEC). **(B)** Chromatogram of final polishing with SEC (blue: absorption at 280 nm, specific for proteins; brown: conductivity; green: buffer gradient). **(C)** SDS-PAGE of PCSK9ΔCTD purification (TEV/TEV2: protein after incubation with different amounts of TEV protease to cleave the ubiquitin tag). **(D)** Chromatogram of final polishing with SEC. A small peak representing free ubiquitin tag is visible and was successfully removed from PCSK9ΔCTD as visible in (C).

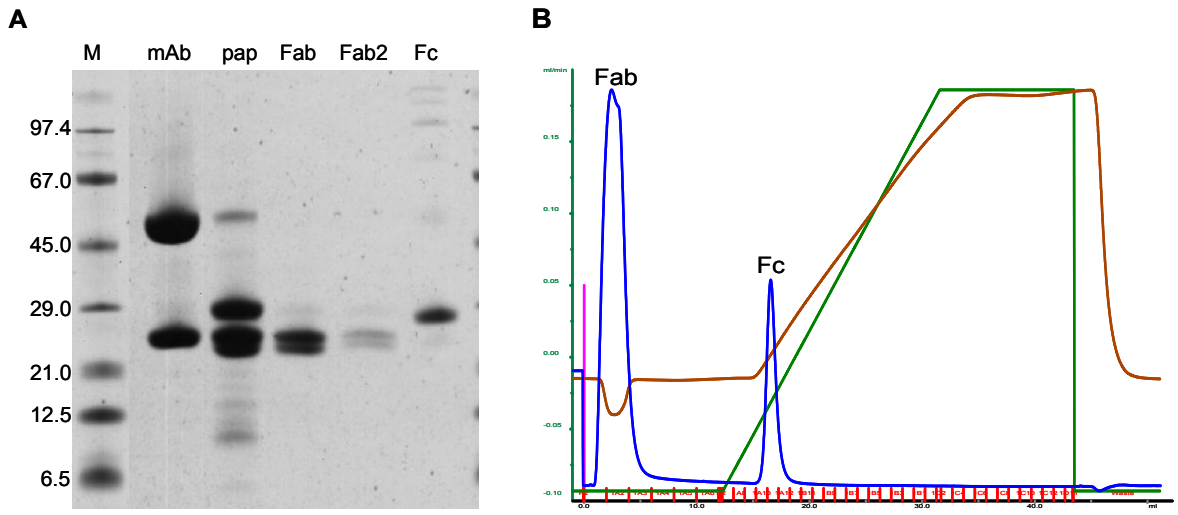


Figure 28: Preparation of Fabs from aPCSK9-mAb1. (A) SDS-PAGE of Fab1 preparation. Light and cleaved heavy chain are visible as distinct bands (mAb: full-length mAb1; pap: sample after papain cleavage; Fab/Fab2: Fabs obtained from protein-A flow-through; Fc: Fc portion from protein-A elution). **(B)** Chromatogram of protein-A affinity purification. Fabs are located in the flow-through, the Fc portion binds to protein-A and is eluted with a pH gradient.

6.2 aPCSK9-Fab1

6.2.1 Expression, Purification and Papain Cleavage of aPCSK9-mAb1

aPCSK9-mAb1 was expressed in hybridomas and purified on MabSelect Xtra (GE healthcare) as described in section 4.3.3. Full-length mAb1 was digested with papain and purified as explained in section 4.3.4 for crystallization purposes. An exemplary SDS-PAGE of the column purification of mAb1and the ensuing papain cleavage and purification of Fab1 is shown in figure 28.

6.2.2 Epitope Identification by Peptide Microarrays

Identification of the linear epitope of aPCSK9-mAb1 was performed by the NBE unit. Custom peptide microarrays (Pepspots) were obtained from JPT (Berlin, Germany) and used to identify the linear epitope to which aPCSK9-mAb1 binds. Based on the deposited PCSK9 X-ray structures, peptide arrays comprising 12 amino acids found on the surface of the protein were designed. These peptide microarrays were blotted with aPCSK9-mAb1, bound mAb was detected with a Cy5-labelled secondary antibody and imaged. The core sequence TPPRYR (representing amino acids Thr162 - Arg167 of PCSK9) was consistently seen in the peptide spots with the strongest binding signals. This sequence may represent only part of the true epitope, since only the incomplete PCSK9 sequence was scanned. In addition, upon binding, antibodies may induce structural changes and expose other a.a. residues with which they may

also interact. Therefore, it was decided to further examine the putative epitope, using biophysical methods and X-ray crystallography as described in the following sections.

6.2.3 aPCSK9-Fab1 Apo Structure

aPCSK9-Fab1 was concentrated to 10 mg/ml prior to crystallization. During concentration the buffer was exchanged against 10 mM TRIS, 20 mM NaCl, pH 7.5. Crystallization trials were set up using the vapor diffusion method by both the hanging and the sitting drop procedures. Equal volumes of protein and crystallization solution were mixed. Uncomplexed aPCSK9-Fab1 crystallized in 0.18 M ammonium sulfate, 27% PEG 4K, 0.02 – 0.14 M copper (II) chloride and 0.1 M TRIS, pH 7.0 – 7.6 after 16 days at 20 °C. The apo crystal structure was solved at a resolution of 2.2 Å. Detailed information on crystallography is given in appendix A.

6.2.4 aPCSK9-Fab1 Complex Structure

aPCSK9-mAb1 was found to bind to a linear peptide on the surface of PCSK9 from Ile161 - Glu170. For complex formation, aPCSK9-Fab1 was incubated with a 1.5 molar excess of peptide ITPPRYRADE (Biotrend) representing this sequence on ice for 1 h and subsequently purified by SEC on a Superdex 75. The complex crystallized in 17.2% PEG 3350, 0.014 M MES, pH 6.0 equilibrated against 24% PEG 3350 after 2 days at 4 °C.The complex structure was solved at a resolution o 1.4 Å. In the complex structure the electron density is very well defined for the binding site region. The three C-terminal amino acids of the peptide (Ala168, Asp169 and Glu170) are disordered in the crystal structure, most probably because the C-terminus is not involved in binding and remains flexible (see figure 29 on the next page). A part of the peptide adopts a conformation similar to a 3_{10}-helix that is defined by H-bonds between a carbonyl oxygen at amino acid position *i* and an amide nitrogen at position *i+3*. In the structure presented here, two H-bonds are formed between Pro163/Tyr166 and Pro164/Arg167. An additional factor for the stabilization of the peptide's conformation is a bifurcated H-bond between the carbonyl oxygen of Pro163 and the guanidine moiety of Arg167. In the conformation seen in the complex structure, the peptide has a solvent accessible surface area of 1035 Å^2 of which 292 and 27 Å^2 are buried by interactions with the light and heavy chain, respectively. All CDRs of the light chain as well as CDR:H2 and CDR:H3 are involved in binding. The peptide's Ile161 is partially defined in the electron density but does not interact with the antibody. Thr162 forms two H-bonds with the side chain hydroxyl and the amide nitrogen to the carbonyl oxygens of L:Ser91 and L:Asn92 in CDR:L3, respectively. In addition, the methyl group of the side chain forms a C-H-π-interaction with L:Trp94 with 3.8 Å

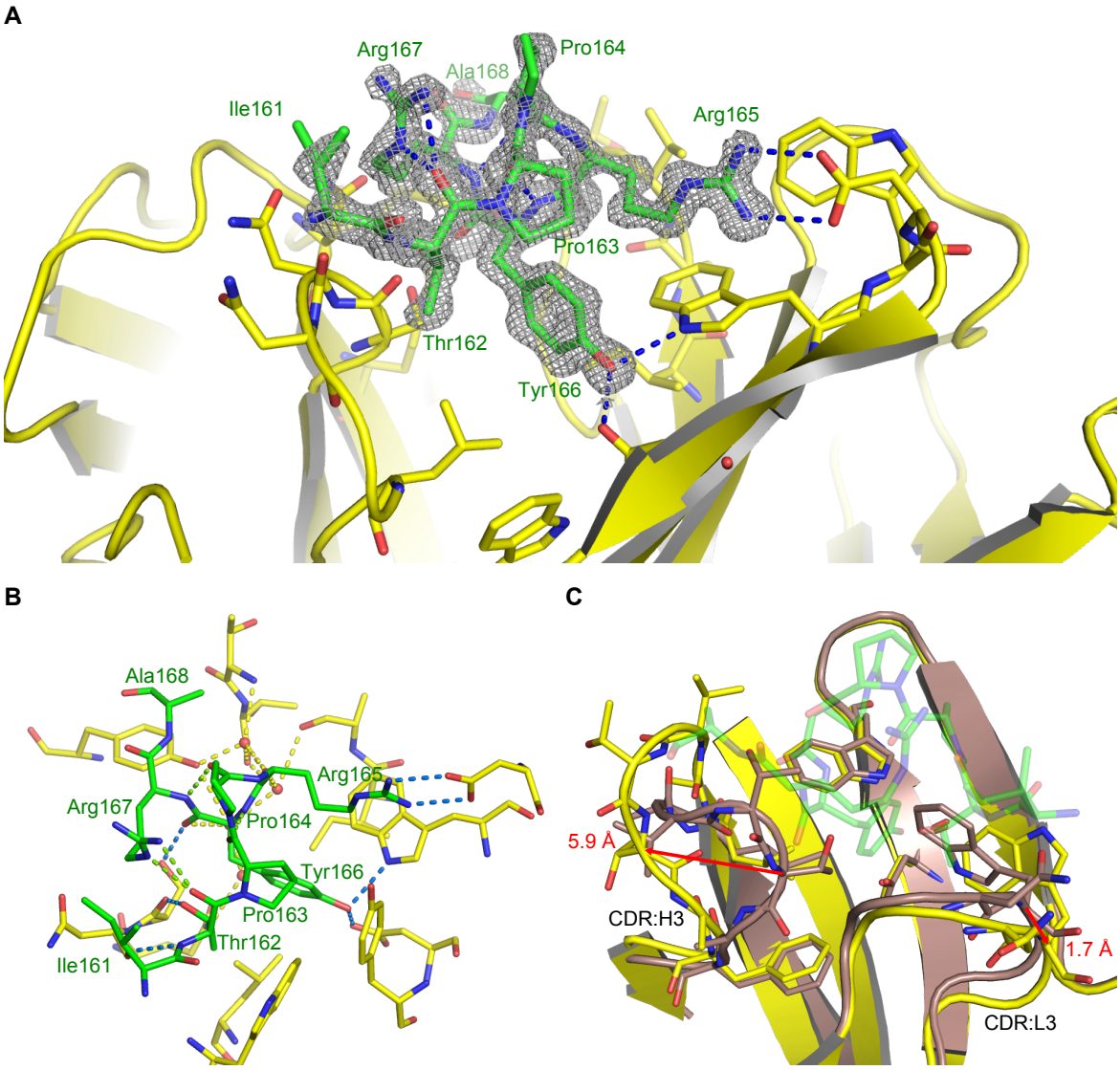


Figure 29: Crystal Structures of aPCSK9-Fab1. (A) Binding site of aPCSK9-Fab1. The electron density for the peptide is shown at a contour level of 1.5 σ and well defined for the majority of residues. The C-terminal residues Asp and Asn are not visible in electron density and do not contribute to binding. Amino acids of aPCSK9-Fab1 within a distance of 3.5 Å to the peptide are shown as sticks. H-bonds are indicated by blue dotted lines. **(B)** Atomic interactions between Fab1 and the decameric PCSK9 peptide. All interacting residues are depicted as sticks. Intermolecular H-bonds are depicted in blue, intramolecular H-bonds in green and H-bonds in a water network in yellow dotted lines. **(C)** Conformational changes in Fab1 upon binding. CDR:L3 increases its distance to the binding site by a maximum backbone movement of 1.7 Å. CDR:H3 shows a maximum distance of backbone c-alpha toms of 5.9 Å. The change in the conformation of both CDRs allows the peptide access to the binding site, especially to the otherwise blocked H:Asp50 and H:Trp52 hot-spot.

as closest distance of non-hydrogen atoms (Brandl et al., 2001). The same interaction is observed between the C4 atom of Pro163 and H:Tyr58 in CDR:H2 (3.7 Å). Arg165 forms a solvent-exposed bidentate salt bridge to H:Asp54 in CDR:H2. The carboxyl oxygens of Arg165 and Tyr166 participate in a H-bond network mediated by several water molecules that are enclosed in a cavity between peptide and Fab1, mainly capped by CDR:H3. The most prominent interaction with aPCSK9-Fab1 is conveyed by Tyr166. It inserts into a cavity framed by light and heavy chain and there forms H-bonds to H:Asp50 and H:Trp52, both located in CDR:H2. These H-bonds contribute significantly to the free energy of binding as they reside in an otherwise hydrophobic vicinity provided by the side chains of L:Trp94, L:Leu96, H:Trp52, H:Tyr58 and H:Ile95. In addition, the phenol ring is bound by T-shaped edge-to-face π-interaction with the side chains of L:Trp94 (3.4 Å) and H:Trp52 (3.5 Å), respectively. Arg167 is the last amino acid for which clear electron density is observed. It does not directly participate in binding to aPCSK9-Fab1, but rather is necessary for intramolecular stabilization of the peptide's conformation, as mentioned above. The structural and affinity data also agree well with findings concerning cross-species reactivity of the antibody. mAb1 binds to cynomolgous PCSK9 where Pro164 is replaced by an Alanine but not to murine PCSK9 that bears more differences in sequence (IIPAWHQTEE) and most notably lacks the tyrosine at position 166.

6.2.5 Conformational Differences Between Apo and Holo Structure

Comparison of the structures of aPCSK9-Fab1 in its apo and complexed form reveal major conformational changes upon binding of the peptide. When both structures are superimposed on the Fv part, these structural rearrangements become obvious (see figure 29 on the facing page). After binding of the peptide, CDR:L3 increases its distance to the binding site by a maximum backbone movement of 1.7 Å (measured from the Cα-atom of L:Trp96) and the side chain conformations of L:Asn93 and L:Trp94 change. In its apo conformation, L:Asn93 would clash with Ile161 of the peptide. The same is true for L:Trp94 as it blocks access to H:Asp50 in the apo structure. The largest conformational change takes place in CDR:H3, more precisely in the heptameric stretch from H:Ile95 to H:Phe100A. The backbone and side chain conformations change dramatically, resulting in a rmsd of 1.9 Å for Cα and 3.2 Å for all atoms, respectively. This effect is most pronounced for H:Thr99, for which the coordinates of the Cα-atom and the hydroxyl oxygen differ by 5.9 Å and 9.6 Å, respectively. In the apo structure, CDR:H3 completely blocks access to the "hot spot" (consisting of H:Asp50 and H:Trp52) that forms H-bonds to Tyr166 which were identified as the most important interaction for binding. As also observed for a aDabiFab2/3

complex (see section 5.3), the conformation seen in the apo structure might be favored in terms of entropy in the absence of a binding partner, as it shields a mostly hydrophobic cavity from contact with the polar water environment.

6.2.6 SPR Peptide Analysis

The X-ray structure presented above reveals the atomic interactions of each amino acid of the linear decameric PCSK9-stretch with aPCSK9-Fab1. To further analyze the individual contributions of each amino acid of the peptide for affinity to aPCSK9-mAb1, SPR alanine scanning was performed. A loop on the surface of PCSK9 was identified as epitope for aPCSK9-mAb1. Based on the structure of PCSK9, three peptides with the amino acid sequences ITPPRYRADE (pep10), TPPRYRAD (pep8) and PPRYRA (pep6) were synthesized and used for an initial affinity determination with SPR. Almost the complete affinity is conveyed via pep8, thus this stretch was used for a more detailed investigation of the contribution of individual amino acids. Each position of the octameric PCSK9-peptide TPPRYRAD from Thr162 – Asp169 was exchanged against alanine and the N- and C-terminal amino acid was deleted, respectively. As the kinetic constants measured for the mutant peptides did not agree well with a 1:1 interaction model, it was chosen to determine the affinity at equilibrium, when a steady-state SPR signal was reached. In case of the high affinity decameric peptide (pep10) and of full-length PCSK9, kinetic analysis was used to determine K_D, as (especially for low concentrations) no steady state signal could be reached within a convenient time period (see figure 30 on the next page). The results of affinity determination are presented and summarized in figure 31 on page 112. The affinity data allows to determine the individual contribution of each amino acid interaction to binding of the peptide. It is obvious that almost the entire affinity of aPCSK9-mAb1 with PCSK9 (K_D = 3.87 nM) is conveyed by a decameric stretch representing the amino acid sequence of PCSK9 from Ile161 to Glu170 (K_D = 4.96 nM). As the octameric peptide TPPRYRAD had an only slightly lower affinity (K_D = 22.1 nM), it was decided to perform the alanine scanning based on the octameric peptide. Upon removal of Thr162, the affinity of the peptide drops 4 orders of magnitude to 170 μM, which can be explained by the numerous interactions formed by this amino acid. Interestingly the hexameric peptide, where in addition to Thr162 also Asp169 was removed, showed a comparable affinity of 77.8 μM. This indicates that Asp169 has no significant contribution to binding and agrees well with the observed affinity of the ΔAsp169-peptide (61.0 nM) and the crystal structure, where this amino acid is disordered. Substitution of Pro163 against alanine decreases the affinity by a factor of 70. In the complex structure Pro163 interacts with H:Tyr58 and, in addition, induces

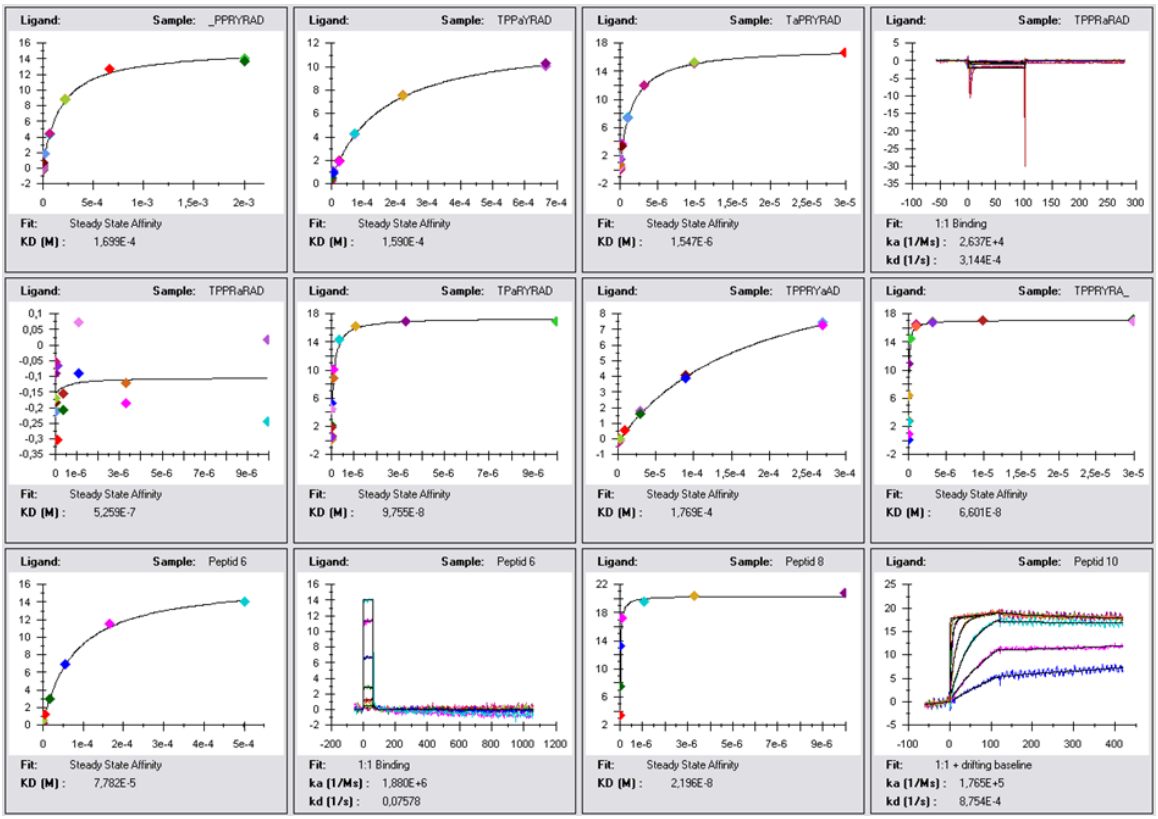
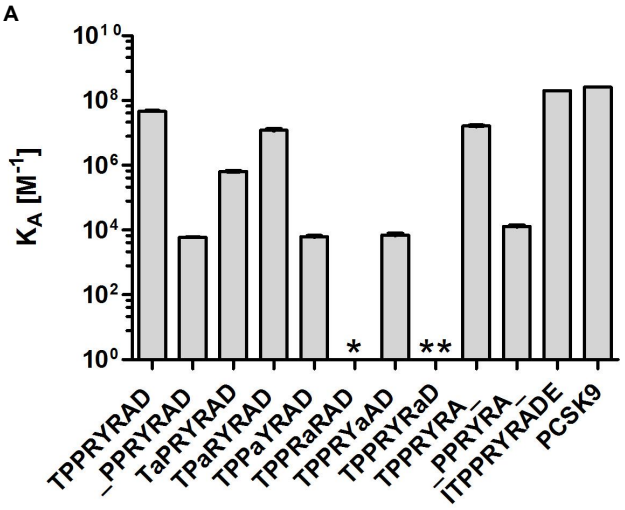


Figure 30: SPR Sensograms of Peptides Binding to aPCSK9-Fab1. Sensograms for the binding of decameric, octameric and hexameric peptides resembling a linear wild-type PCSK9 stretch are given as well as mutants thereof. All sensograms were evaluated at equilibrium concentrations. Exemplary sensograms are shown for peptide 6 and mutant TPPRaRAD. The mutant shows no binding to aPCSK9-Fab1. Binding of peptide 10 was analyzed by kinetic rate constants as no equilibrium was reached in a convenient period of time.



B

Peptide	K_D [M]	%CV
TPPRYRAD	2.20E-08	8.7
_PPRYRAD	1.70E-04	4.9
TaPRYRAD	1.55E-06	1.7
TPaRYRAD	8.32E-08	8.1
TPPaYRAD	1.59E-04	3.5
TPPRaRAD	*	
TPPRYaAD	1.43E-04	11.2
TPPRYRaD	**	
TPPRYRA_	6.10E-08	2.1
PPRYRA	7.78E-05	7.6
ITPPRYRADE	4.96E-09	n.d.
PCSK9	3.87E-09	n.d.

Figure 31: Affinity of aPCSK9-Fab1 for PCSK9 Peptides. (A) Graphical representation of the affinities of full-length PCSK9 and peptides thereof. **(B)** Table presenting K_D and the respective coefficient of variation (* no binding was observed, ** no alanine substitution possible).

the formation of a short 3_{10}-helix. Only in this helix conformation the observed interactions of the peptide with Fab1 are possible. Proline is frequently observed at the N-terminus of helices (Richardson and Richardson, 1988). With its restricted backbone conformation (Φ value of $\sim 60°$) it is close to the theoretical backbone Φ value of a 3_{10}-helix (Venkatraman et al., 2001). As a consequence of the closed pyrrolidine ring, it possess only one rotatable angle and thus loses less entropy than other amino acids in forming a secondary structure. In addition there is no unsatisfied H-bond to proline's nitrogen atom as it is covalently bound in the ring. Two prolines have been shown to induce 3_{10}-helix helices in short peptides (Hanessian et al., 2004). Formation of the 3_{10}-helix might thus be disadvantageous upon removal of Pro163. In contrast to this, mutagenesis of Pro164 results in an only 4-fold affinity decrease. Pro164 does not participate in interactions with the antibody as revealed by the crystal structure. It forms only one H-bond in the 3_{10}-helix, but this may also be formed by the alanine in case of the mutant peptide. Furthermore, Richardson has shown that prolines occur more often at the first than at the second position of these helices (Richardson and Richardson, 1988). When Arg165 and Arg167 are substituted, the K_D reduces to 159 μM and 143 μM, respectively. The salt bridge formed between Arg165 and H:Asp54 and the H-bonds formed by Arg167 with the peptide backbone that stabilize the peptide's conformation, thus seem to contribute equally to binding affinity. The most important interaction is mediated by Tyr166, as it forms two H-bonds to H:Asp50 and H:Trp52 and van der Waals interactions in an otherwise hydrophobic environment. As a consequence, when Tyr166 is substituted against alanine, no binding is observed in SPR.

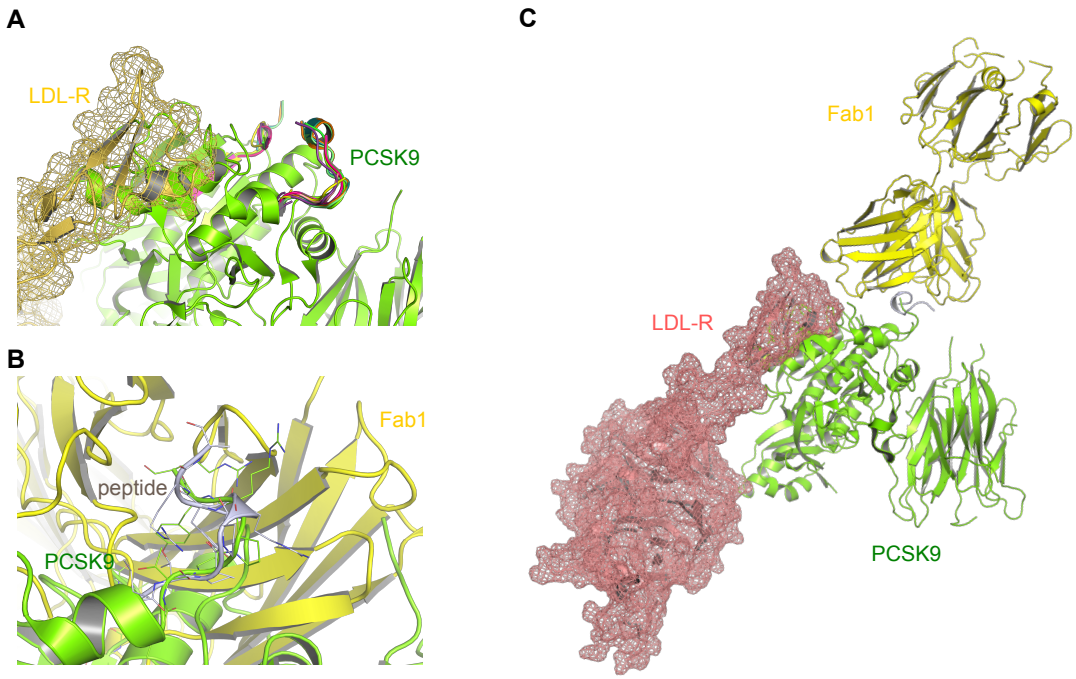


Figure 32: Modeled Complex of aPCSK9-Fab1:PCSK9. (A) Superposition of the linear stretch to which mAb1 binds from all PCSK9 structures deposisted. The stretch has different conformations and electron density is missing, which indicates the intrinsic flexibility of this loop. (PCSK9 colored in green, LDL-R yellow) **(B)** Superposition of the Fab1:peptide structure with PCSK9 on the linear stretch. The peptide (colored in silver) adopts the similar conformation as seen in full-length PCSK9, but Fab1 massively clashes with PCSK9. **(C)** Fab1 rotated around the peptide to minimize possible clashes with PCSK9. The final model no longer interferes with binding of LDL-R (red).

6.2.7 aPCSK9-mAb1 Does Not Block EGF-A Binding

aPCSK9-mAb1 was identified as functional in an alpha screen blocking PCSK9 interaction with the LDLR and also to increase LDL uptake in a cellular assay. With peptide chip analysis it was discovered that mAb1 binds a linear epitope on PCSK9 that is close to the LDLR binding site. Based on the crystal structure of Fab1 in complex with the decameric PCSK9 peptide it was aimed to model the complex of Fab1 with full-length PCSK9 (see figure 32). Analysis of all PCSK9 structures available in the pdb (30. January 2013) was performed to check, whether the peptide in the X-ray structure represents a conformation that has already been observed. In fact the stretch from Thr162 – Tyr166 can be superimposed to one of the deposited structures (PDB ID 3BPS) with a rmsd of 0.73 Å (over the 5 Cα-atoms) where Pro163 also interacts with Tyr166 to induce a β-turn (Kwon et al., 2008). When the complex of aPCSK9-Fab1 and the peptide is superimposed this way, there is a variety of clashes between Fab1 and PCSK9. In numerous structures (Bottomley et al., 2009; Chan et al., 2009; Cunningham et al., 2007; Hampton et al., 2007; Li et al., 2003;

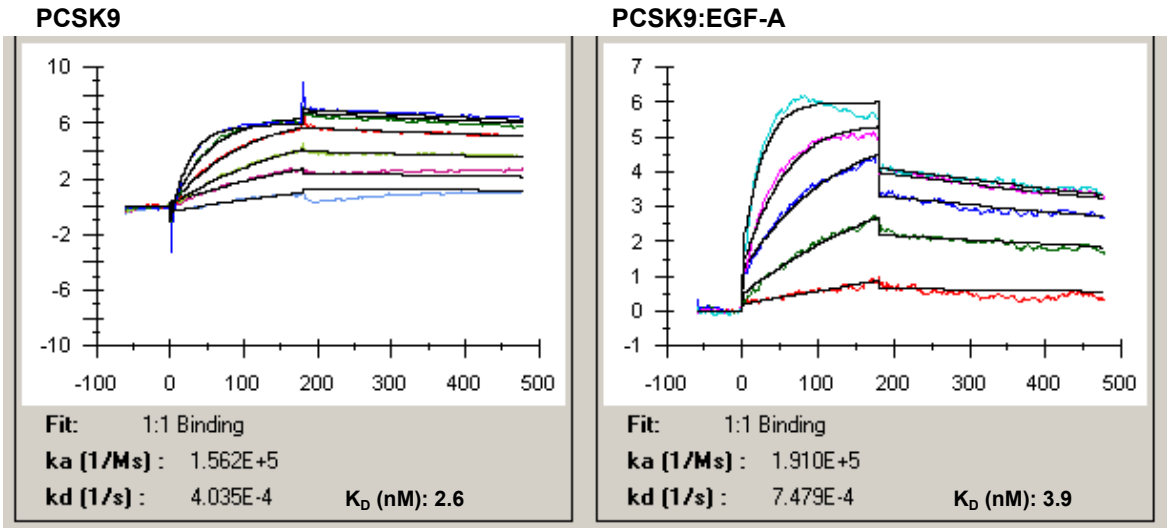


Figure 33: SPR Sensograms of Binding of PCSK9 and PCSK9:EGF-A to mAb1. Increasing concentrations of PCSK9 and PCSK9 complexed to EGF-A were injected. Data fitted to a 1:1 binding model (black lines). Both proteins bind to mAb1 with comparable affinity.

Liang et al., 2012) the loop in PCSK9 resembling the peptide is defined only until Arg165 (PDB-IDs 2W2O, 2W2M) or Arg167 (3H42, 2QTW, 3M0C) or Ala168 (2P4E, 3SQO), respectively and the first residue defined by electron density C-terminal to this gap is Gly176. Thus, although the conformation of the decameric peptide differs from the previously observed conformations in full-length PCSK9, it is sound to assume this stretch is highly flexible and the formation of the short 3_{10}-helix might be induced by the antibody. As a consequence aPCSK9-Fab1 was rotated around the peptide (while the position of the N-terminal amino acids was fixed to those seen in full-length PCSK9) to obtain a minimum number of clashes with full-length PCSK9. This resulted in a complex model where Fab1 does not block the EGF-A binding site and does not interfere with binding of PCSK9 to EGF-A. Because of this ambiguous model, a SPR analysis was performed where mAb1 was captured on a CM5 chip and PCSK9 or PCSK9:EGF-A complexes were injected in increasing concentrations (see figure 33). It was found that EGF-A did not affect binding of PCSK9 to mAb1, both bind with an affinity of 2.6 and 3.9 nM, respectively. As the LDLR protein is much larger than EGF-A, it is still possible that mAb1 reduces the affinity of the receptor by interfering with interactions of, for example, the ligand binding domains (LBDs) with PCSK9. LBDs are thought to interact with the CTD of PCSK9. As mAb1 binds adjacent to the N-terminus of EGF-A it might occupy space that is otherwise needed for proper rearrangement of the LBDs. This model offers a possible explanation why mAb1 is capable of increasing the LDL-C uptake in a cellular assay, although it does not inhibit binding of EGF-A to PCSK9.

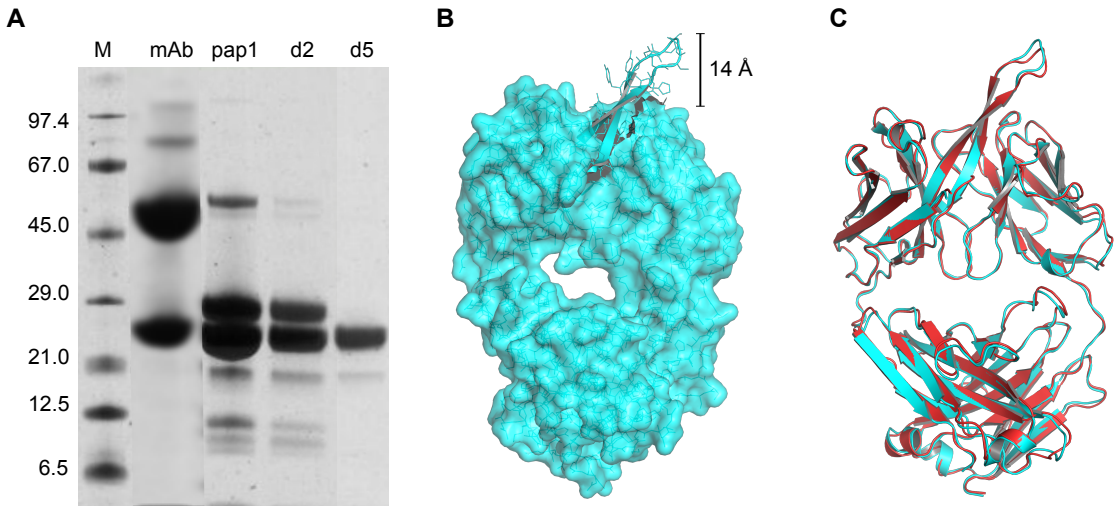


Figure 34: Preparation of aPCSK9-Fab2 by Papain Cleavage and Structural Features of the Fab. (A) Fc portion and uncleaved heavy chain were completely digested after incubation for 5 days at 8 °C (mAb: full-length mAb1 (lane is pasted electronically for the sake of clarity); pap: sample after papain cleavage; d2/5: sample after 2 and 5 days of incubation at 8 °C) **(B)** Fab2 has a long CDR:H3 loop comprising 20 amino acids that protrudes more than 14 Å from the flat surface of the paratope. **(C)** Except of a small 1.5 Å backbone movement in the apex of CDR:H3, there are no conformational differences between free (cyan) and complexed (red) Fab2.

6.3 aPCSK9-Fab2

6.3.1 Expression, Purification and Papain Cleavage of aPCSK9-mAb2

aPCSK9-mab2 was expressed in CHOs and purified on MabSelect Xtra (GE healthcare) as described in section 4.3.3. Full-length mAb2 was digested with papain as explained in section 4.3.4 for crystallization purposes. Surprisingly the Fc portion could not be separated from the Fab by protein-A affinity chromatography. By coincidence it was found, that incubation of the Fab-Fc mixture after papain digestion for 5 days at 8 °C resulted in complete proteolysis of the Fc portion and uncleaved heavy chain. This is shown with an exemplary SDS-PAGE of papain cleavage and incubation in figure 34. This may be caused by unspecific cleavage either by papain detached from the beads during incubation or by endogenous proteases from CHOs that are purified in trace amounts during initial protein-A affinity chromatography.

6.3.2 SPR Analysis

Binding to CTD The affinity of aPCSK9-mAb2 to PCSK9 was determined by injecting increasing concentrations of PCSK9 and PCSK9ΔCTD over a SPR chip with mAb2 captured. PCSK9 binds with a K_D of 4.03 nM (k_a = 3.45E5 $M^{-1}s^{-1}$, k_d = 1.39E^{-3} s^{-1}). For PCSK9ΔCTD no specific binding was observed, so the SPR data suggested that

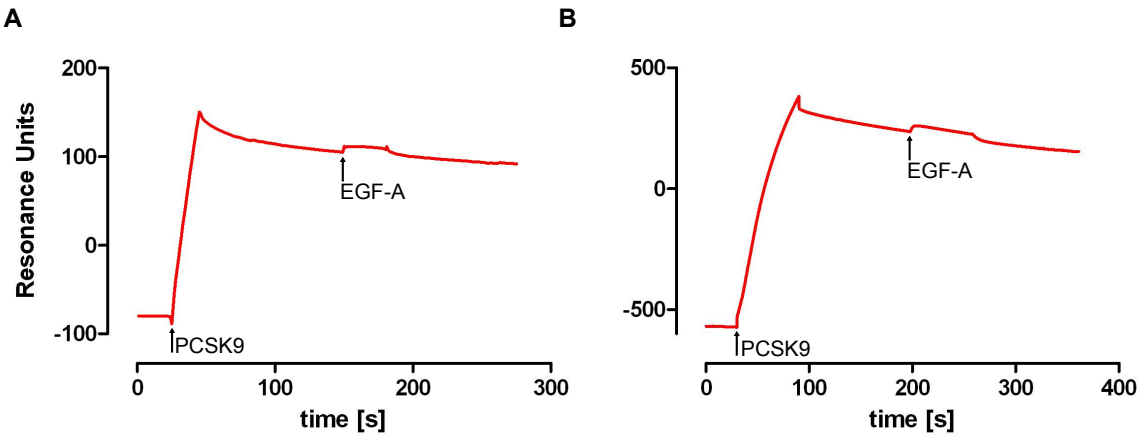


Figure 35: mAb2 Does Not Affect Binding of EGF-A PCSK9. (A) PCSK9 with a C-terminal His-tag is captured by an anti-His antibody immobilized on a CM5 chip. EGF-A is subsequently injected. **(B)** PCSK9 is captured by aDabi-mAb2 immobilized on a CM5 chip. EGF-A is subsequently injected. and still able to bind PCSK9. mAb2 does thus not interfere with binding of EGF-A to PCSK9.

mAb2 binds to the C-terminal domain of PCSK9.

Inhibition of Interaction of PCSK9 With EGF-A EGF-A binds to PCSK9 with a K_D of 1.5 μM as determined by SPR. To investigate whether aPCSK9-mAb2 is able to inhibit interaction of EGF-A with PCSK9 two experiments were performed. PCSK9 was captured with an antibody directed against its C-terminal his-tag. Then EGF-A with a concentration of 100 μM was injected. In a second experiment PCSK9 was captured with mAb2 and EGF-A was injected at the same concentration (see figure 35). The presence of mAb2 did not affect binding of EGF-A to PCSK9. This supports the finding that mAb2 binds to the CTD of PCSK9 that is distant to the EGF-A binding site, as described previously.

6.3.3 Apo Structure of aDabi-Fab2

To further investigate the binding mode, crystallization trials of aPCSK9-Fab2 in its apo form and in complex with PCSK9 were set up. The apo form of aDabi-Fab2 crystallized in 12.6% PEG 3350, 0.01 M spermine tetrahydrochloride, 0.0126 M trisodium citrate (pH 5.0) equilibrated against 24% PEG 3350 after 7 days at 20 °C. The structure was solved at 2.2 Å. The electron density for the whole model is well defined but is becoming weaker towards the apex of CDR:H3, indicating a certain flexibility for this loop. Interestingly CDR:H3 is a very long loop consisting of 20 amino acids which is more often observed in cameloid VHH antibodies and rarely in IgGs (Muyldermans et al., 1994). In humans and mice the average length for CDR:H3 is around 13 - 14 amino acids (Ivanov et al., 2005) which is significantly

shorter than in aPCSK9-mAb2. As a result CDR:H3 protrudes from the otherwise flat paratope by more than 14 Å, a feature rarely seen for antibody structures (see figure 34 on page 115). As the binary complex of Fab2 and PCSK9 did not crystallize, several ternary complexes consisting of PCSK9, Fab2 and additional Fabs were tested for crystallization. A ternary complex with Fab3H42 finally yielded diffraction quality crystals.

6.3.4 Expression and Purification of Fab3H42

mAb3H42 was expressed in HEK293 cells and purified on a protein-A affinity column. The ensuing digestion by papain and separation of Fabs and Fc portion on a protein-A column were performed as described in section 6.2.1. The Fab was exclusively used for crystallization of a ternary complex consisting of aPCSK9-Fab2, PCSK9 and Fab3H42.

6.3.5 Ternary Complex of aDabi-Fab2

As no crystals formed for the binary complex of Fab2 and PCSK9, screening for crystallization of a ternary complex consisting of Fab2, PCSK9 and an additional Fab for which the complex structure with PCSK9 had previously been reported (Chan et al., 2009), was performed. First, needle-shaped crystals appeared in 0.2 M sodium acetate, 16% PEG 4K and 0.1 M TRIS, pH 8.5 but they did not diffract. After extensive rounds of finescreening and microseed matrix screening (Ireton and Stoddard, 2004), it was possible to reproduce plate-shaped crystals in 0.7 M lithium sulfate and 14.5% PEG 8K that diffracted to a maximum resolution of 3.2 Å at a synchrotron beamline. With spacegroup *C*2 and cell constants of a=262.19 Å, b=138.65 Å, c=69.54 Å and β=102.87°, the ternary complex displays exactly the same metric as the binary complex with PDB-ID 3H42. As it turned out, Fab2 occupies solvent channels that were present in the binary complex reported (see figure 36 on the next page). The electron density is well defined for Fab3H42 and PCSK9, but is weaker for Fab2, especially with increasing distance from the binding site. Accordingly the average B-factor is very high for Fab2 (148 Å^2) compared to Fab3H42 (42 Å^2) and PCSK9 (60 Å^2). Fabs are known to be flexible at the linker of constant and variable domains. TLS parameterization was thus performed with one TLS group per domain of Fab2. Analysis of the anisotropic thermal ellipsoids of Fab2 showed that the constant domains of the Fab have the highest B-factor. In the X-ray structure, the constant domains of Fab2 are involved in only 2 water mediated crystal contacts which might allow for domain movements that could explain the increased flexibility and weak electron density compared to the variable domains that are bound to the

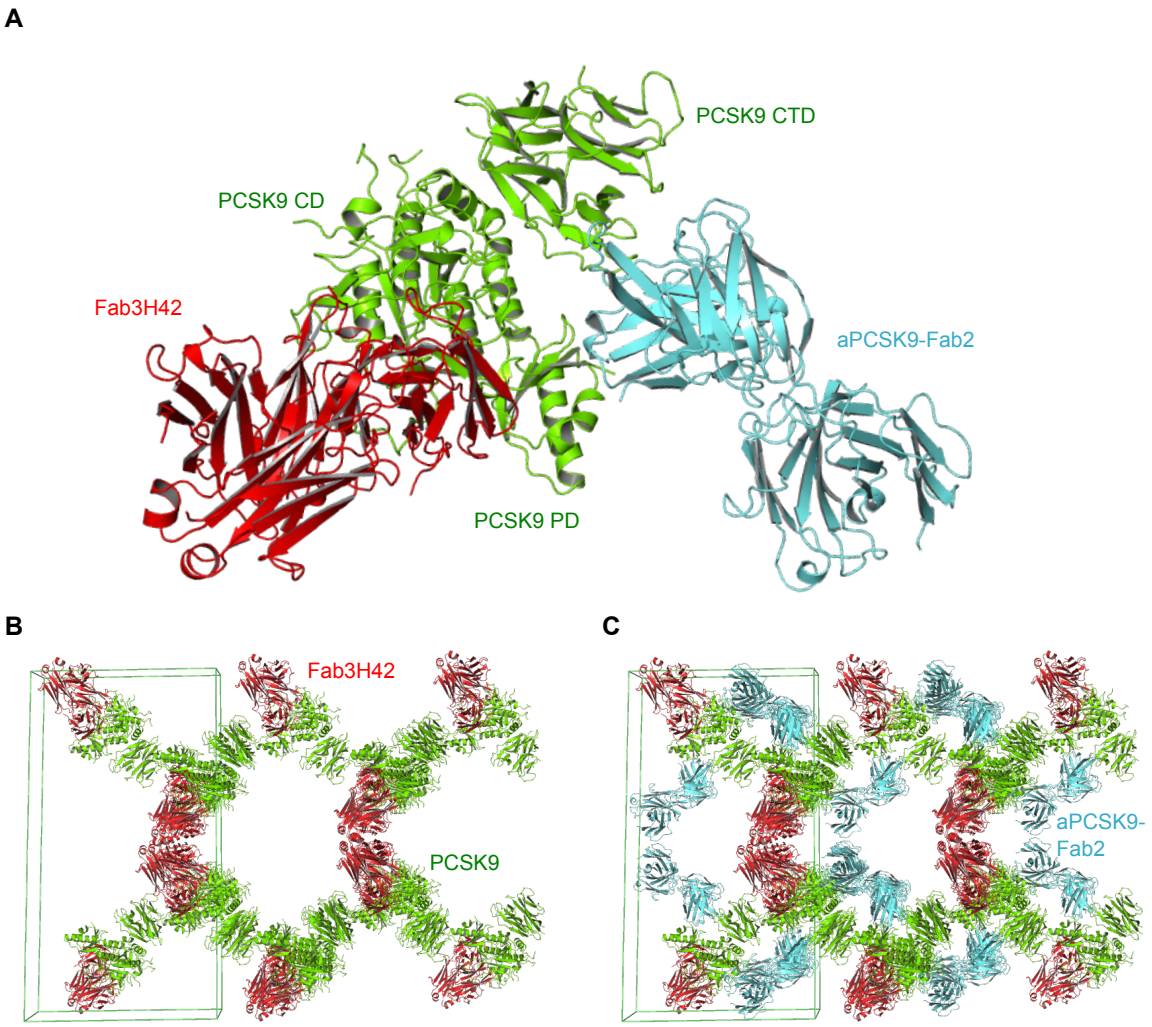


Figure 36: Ternary Complex of aDabi-Fab2. (A) The crystal structure of a ternary complex consisting of aPCSK9-Fab2, PCSK9 and Fab3H42 was solved at 3.2 Å resolution. aPCSK9-Fab2 binds to module M1 of the C-terminal domain of PCSK9. **(B)** Crystal environment of the binary complex PCSK9:Fab3H42 (PDB-ID 3H42), view along the b-axis. **(C)** Crystal environment of the ternary complex. The cell metrics are the same as in the binary crystal. aPCSK9-Fab2 occupies the large solvent channels present in (B).

CTD of PCSK9. Interestingly the B-factors steadily decrease when moving from the constant to the variable domains of Fab2 and finally to the CTD of PCSK9, although independent TLS groups were chosen for each domain. In addition, analysis of the anisotropic B-factors indicates flexibility longitudinal to the b-axis of the crystal that coincides with the direction of continuous solvent channels. This makes sense as a transversal movement is restricted by the neighboring molecules. aPCSK9-Fab2 binds mostly to a convex surface in the CTD of PCSK9 with a contact area of 794 Å^2 to which the heavy and light chain contribute 554 Å^2 and 240 Å^2, respectively. All CDRs of the heavy chain as well as CDR:L3, participate in binding to PCSK9. There are no major conformational changes upon binding in the structures of Fab2 and PCSK9, compared to their apo conformations. Only the apex of CDR:H3 slightly moves by a maximum of 1.5 Å but there are no sidechain reorientations (see figure 34 on page 115). The C-terminal domain of PCSK9 is cylindrically shaped and has a resistin-like fold with quasi 3-fold symmetry and each of the three modules comprises a β-sandwich made of 6 β-strands that is stabilized by 3 disulfide bonds (Cunningham et al., 2007; Piper et al., 2007). Within the CTD, Fab2 binds exclusively to module 1 (M1) and forms numerous H-bonds and hydrophobic interactions. Fab2 contacts 2 loop regions from Ser465 – Thr473 and from Phe515 – Glu518 and also interacts with Arg496 and Arg510. There is an additional contact area of 120 Å^2 between H:Ala74 and H:Lys75 in the framework region 3 of the heavy chain (FR:H3) and the N-terminal stretch of the prodomain of PCSK9 from Thr61 – His65. SPR data suggests that the main contribution for binding and selectivity is likely to occur via the CTD and the costructure confirms this assumption. The interaction of of the Fab's FR:H3 with the prodomain of PCSK9 seems to be a secondary effect that takes place only after specific binding of the antibody's paratope to the CTD.

6.3.6 Mode of Action of aPCSK9-mAb2

After superimposing the aDabi-Fab2 complex with the structure of PCSK9 in complex with EGF-A or LDLR (see figure 37 on page 121), it becomes obvious that binding of Fab2 neither blocks access of PCSK9 to EGF-A nor to LDLR (Kwon et al., 2008; Rudenko et al., 2002; Surdo et al., 2011). Fab2 does also not inhibit the interaction of the b-propeller domain of LDLR with the prodomain of PCSK9 (Kwon et al., 2008). So the LDL-lowering effect of Fab2 must be caused by binding to the CTD while not blocking the interaction of PCSK9 with LDLR at neutral pH. The CTD was shown to contribute to the reduction of LDLR on cell surface, while the underlying mechanism is not well understood (Zhang et al., 2008). In a cellular LDL uptake assay, PCSK9ΔCTD displayed a 20 fold higher IC50 compared to the wild type (Ni et al.,

2010) while there are no differences in affinity for LDLR (Bottomley et al., 2009). Interestingly, Fab2 binds to an epitope of PCSK9 in which several mutations correlating with the LDL-level have been identified. It directly binds Arg469, Arg496 and Phe515. In addition, access to Ile474 is sterically blocked by Fab2 so that interaction of this amino acid with another protein seems unlikely (see figure 37 on the facing page). Individuals with the natural occurring mutations R469W, R496W or F515L display an ADH-phenotype (Allard et al., 2005; Kotowski et al., 2006; Pisciotta et al., 2006). The variant I474V is not completely understood, as it has been associated with both high and low levels of LDL-C (Abifadel et al., 2003). Besides the genetic evidence for the functional significance of this region, the CTD contains several flexible regions that may be involved in binding to other proteins and putative cofactors (Cunningham et al., 2007; Hampton et al., 2007; Piper et al., 2007). mAb2 may thus act by mimicry or inhibition of putative additional proteins. A possible protein that might be affected by mAb2, is annexin A2 (AnxA2). It has been identified to bind to the CTD of PCSK9 with its N-terminal repeat R1 (Mayer et al., 2008). Notably, AnxA2 was shown to inhibit PCSK9-directed LDLR degradation in a cellular assay and AnxA2-derived peptides blocked PCSK9 LDLR-interaction *in vitro*. Overexpression of AnxA2 in mouse liver increased LDLR levels *in vivo* (Seidah et al., 2012). mAb2 might mimic the effect of AnxA2 by binding to the same site of PCSK9, albeit this is not necessary to inhibit LDLR internalization. In a recent study an antibody against the CTD was identified which reduced 50% of the PCSK9-dependent inhibitory effects on LDL uptake in a cellular assay, while not competing with the binding of AnxA2 (Ni et al., 2010). mAb2 as well as the previously described CTD-binding antibody, both result in partial inhibition of PCSK9 effects in a cellular assay, while antibodies that block the EGF-A binding site of PCSK9 are capable of inhibiting the effects of PCSK9 completely (Chan et al., 2009; Stein et al., 2012). Alternatively, it may also be possible that mAb2 interferes with the binding of PCSK9 CTD to the ligand-binding domain (LBD) repeats of LDLR. Yamamoto et al. proposed a two-step binding model: Following initial interaction of PCSK9 with the EGF-A domain and internalization of the PCSK9:LDLR complex, binding between CTD and LBDs takes place at endosomal low pH (Yamamoto et al., 2011). It was suggested that the CTD gets positively charged by protonation of its numerous surface-exposed histidine residues at endosomal pH and thereby promotes interaction with negatively charged elements in the LBD, although with a series of LDLR deletion mutants it was also shown that no specific LBD-repeat is essential (Zhang et al., 2008). Lo Surdo and colleagues recently reported the crystal structure of PCSK9 in complex with LDLR at neutral pH (Surdo et al., 2011). In this structure, PCSK9 constrains the LDLR in an extended conformation and thereby inhibits domain rearrangements to a closed conformation that are

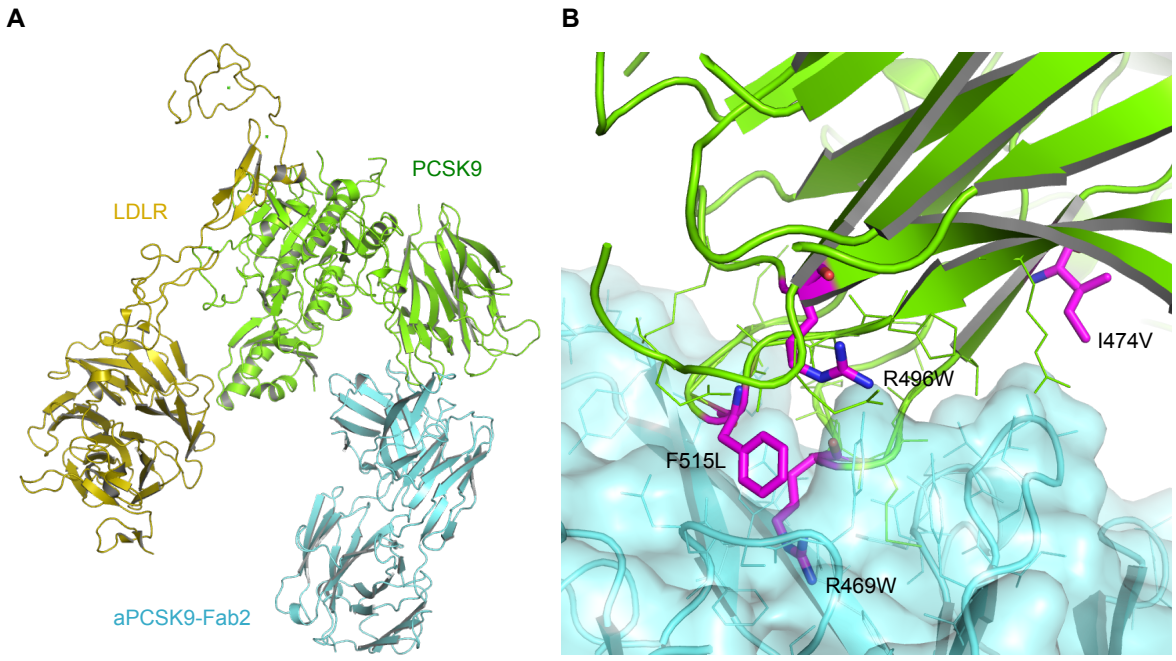


Figure 37: Mode of Action of aPCSK9-Fab2. (A) Superposition of aPCSK9-Fab2 bound to PCSK9 with the PCSK9:LDLR complex (PDB ID 2P5B). Fab2 mainly binds to the C-terminal domain of PCSK9 and has a secondary interaction site with the prodomain. Binding of Fab2 does not interfere with the interaction of PCSK9 to LDL receptor (mAb2 is colored in pale cyan, PCSK9 in green and LDL-R in yellow). **(B)** Close-Up of the antigen binding interface. Fab2 directly interacts with Arg469, Arg496 and Phe515 and sterically blocks access of putative cofactors to Ile474 (mentioned amino acids depicted as pink sticks). All of the highlighted amino acids have been observed in naturally occurring loss- or gain-of-function mutants. mAb2 therefore acts by blocking a region that has been shown to be important for LDLR degradation by genetic findings. This seems to be sufficient for partial lowering of LDL-C levels *in vivo*.

necessary for recycling of the receptor either through direct or allosteric competition (Rudenko et al., 2002; Zhao and Michaely, 2008). LBDs L1 – L6 are flexible in the crystal at neutral pH (thus not defined in electron density) and molecular modeling indicated that even in the extended conformation an interaction of LBDs with the CTD is possible. This putative interaction might be inhibited by mAb2. Although the exact mode of action of aPCSK9-mAb2 needs further investigation, this is the first antibody directed against the CTD of PCSK9 that is able to partially reduce LDL-C levels *in vivo* in nonhuman primates. At the moment there are promising results for antibodies that directly block EGF-A interaction with PCSK9. They effectively lower LDL-C levels and can be used as addon therapy to statins (Sullivan et al., 2012; Stein et al., 2012). Nevertheless, in patients were only a partial reduction of PCSK9 effects is desired, CTD-binding Antibodies like aPCSK9-Fab2 may be superior to antibodies blocking the EGF-A binding site.

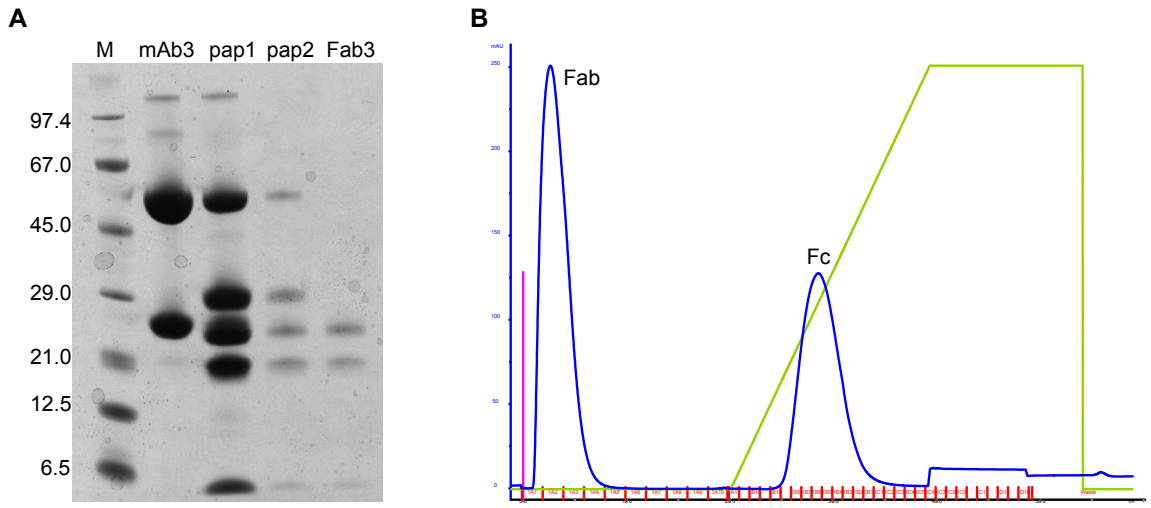


Figure 38: Preparation of Fabs from aPCSK9-mAb3. (A) SDS-PAGE of Fab3 preparation. Light and cleaved heavy chain are visible as distinct bands (mAb3: full-length mAb3; pap1/2: samples after papain cleavage; Fab3: Fabs obtained from protein-A flow-through). A band below 6.5 kDa indicates additional cleavage of the light chain. **(B)** Chromatogram of protein-A affinity purification. Fabs are located in the flow-through, the Fc portion binds to protein-A and is eluted with a pH gradient.

6.4 aPCSK9-Fab3

6.4.1 Expression, Purification and Papain Cleavage of aPCSK9-mAb3

aPCSK9-mab3 was identified to inhibit LDLR interaction with PCSK9 *in vitro* and PCSK9 function in a cellular LDL uptake assay. mAb3 was expressed in CHOs and purified on MabSelect Xtra (GE healthcare) as described in section 4.3.3. Full-length mAb3 was digested with papain and purified as explained in section 4.3.4 for crystallization purposes. An exemplary SDS-PAGE of purified mAb3 and the ensuing papain cleavage and purification of Fab3 is shown in figure 38. As mAb3 obviously contained two additional cleavage sites in the $V_L\varkappa$ domain, Fab3 was directly expressed in HEK293 cells and also isolated and refolded from *E. coli* inclusion bodies. These antibody fragments were of high purity and bound to PCSK9 with the same affinity and kinetic rate constants as full-length mAb3, but failed to crystallize either alone or in various complexes with PCSK9 and other Fabs. As the antibody did also not display efficacy *in vivo*, it was not longer worked on and is described here for the sake of completeness.

6.4.2 Crystal Structure of aPCSK9-Fab3

As mentioned above, aPCSK9-Fab3 failed to crystallized in various ternary complexes with PCSK9 as well as in binary complexes with PCSK9 or PCSK9ΔCTD. There were no successful cocrystallizations with either the Fabs refolded from *E. coli*, Fabs ex-

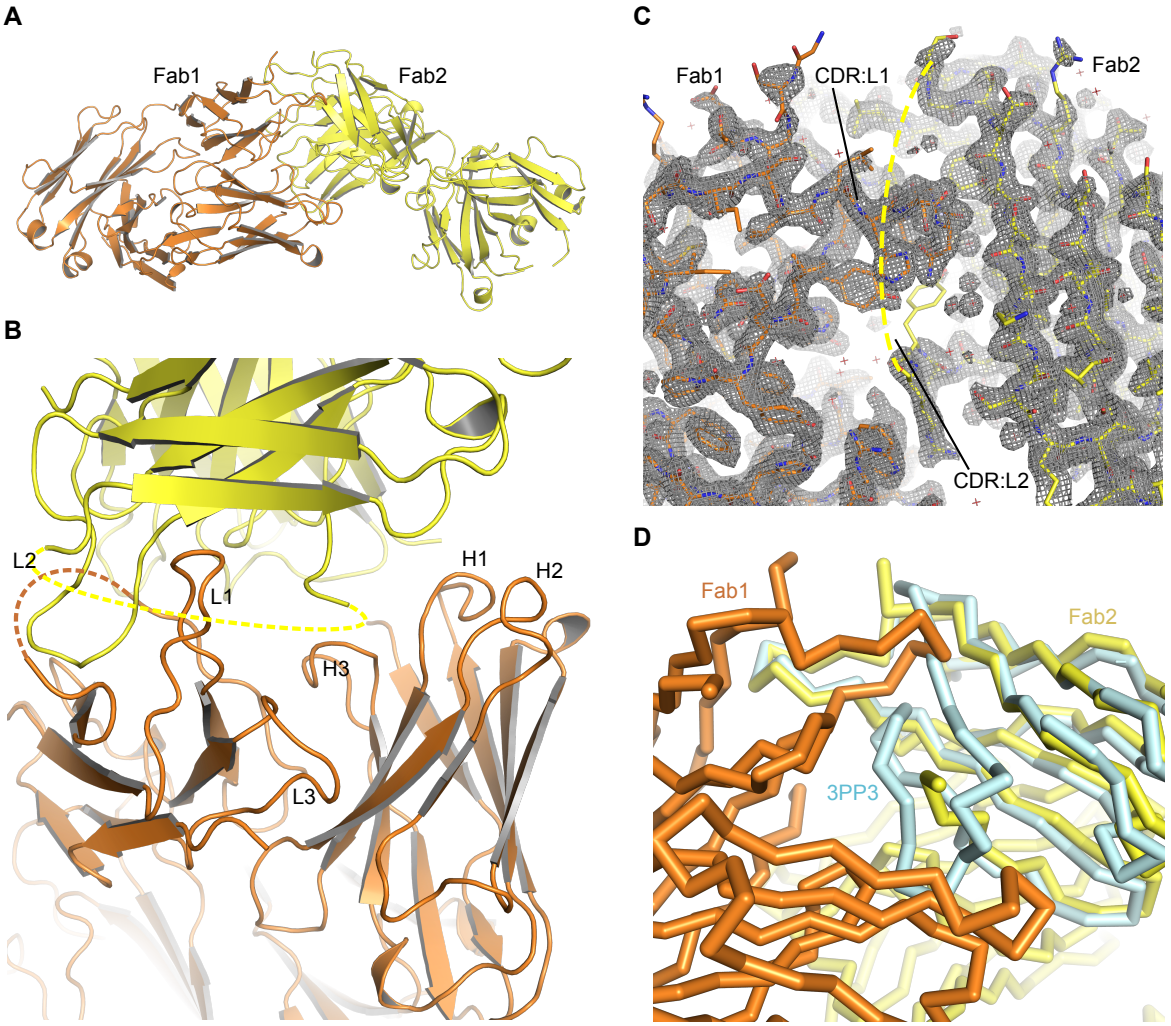


Figure 39: X-ray Structure of aDabi-Fab3. (A) Two protomers are present in the asymmetric unit of the crystal. Their paratopes form interactions that would not be possible for intact Fabs. **(B)** The interface of the two protomers is shown. The missing loop for CDR:L2 is indicated as dotted lines. **(C)** Electron density at the interface of Fab1 and Fab2. Density clearly shows the presence of CDR:L1 in an area that would otherwise be occupied by CDR:L2 in an intact Fab. **(D)** Ribbon representation of the binding interface superpositioned with a homologous Fab (PDB-ID 3PP3). The intact Fab (cyan) would clash with the interaction observed in the crystal.

pressed in HEK293 cells or the Fabs obtained from papain cleavage of mAb3. Only Fab3 from papain cleavage crystallized in its apo form in 15% PEG 3350, 0.0126 M tri-sodium citrate (pH 5.5) equilibrated against 24% PEG 3350 after 7 days at 20 °C. The crystal structure was solved at a resolution of 1.95 Å containing two protomers in the ASU and is presented in figure 39 on the preceding page. Both protomers display virtually identical backbone and sidechain conformations but differ in their elbow angles (141.1° vs. 148.8°) . The electron density is missing for the a stretch from L:Met57 - L:Ser71 that includes the complete CDR:L2. As the SDS-Gel shown in figure 38 on page 122 already indicates, aPCSK9-Fab3 has two additional papain cleavage sites. During incubation with papain, a stretch in the $V_{\kappa}L$ domain is cleaved off and dissociates from the Fab. Increased flexibility of this stretch could also be a cause for missing electron density but this can be ruled out, as here is a third band visible in the SDS gel below 6.5 kDa that represents the cleaved stretch. Superposition of the $V_{\kappa}L$ domain of a deposited Fab structure (PDB ID 3PP3) with high structural homology offers an explanation why the intact Fabs failed to crystallize: The majority of intermolecular interactions between the two protomers in the ASU of the crystal is formed by the two light chains and would not be possible in intact Fabs. Interestingly the interactions seen are only possible in the absence of the cleaved stretch. The CDR:L1 of one protomer is located in the exact area that would otherwise be occupied by CDR:L2 of the second protomer. Intermolecular H-bonds between the light chains are formed by the side chains of L:Asn36, L:Thr37, L:Trp41, L:Tyr55 and L:Tyr100. The intermolecular backbone H-bonds between L:Asn34 and L:Thr78 could also not be formed by intact Fabs. In addition, the part from L:Leu54 - L:Arg56, N-terminal to the missing stretch, changes its conformation (that is similar in all intact antibodies) and blocks access to the binding site of the NCS-related Fab. Especially L:Arg56 resides on top of the paratope and forms H-bonds to the sidechain of H:Lys32 and the backbone carbonyl oxygen of H:Glu102. In fact the only intermolecular H-bonds between the heavy chains are formed by H:Tyr33 and H:Arg106.

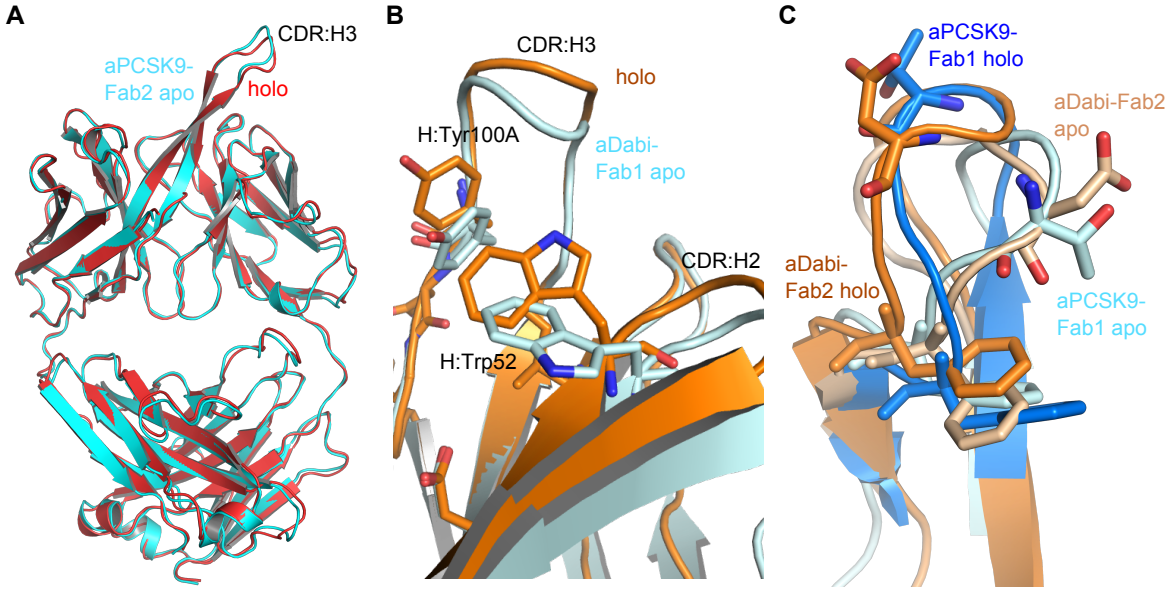


Figure 40: Conformational Differences of the Paratope Between Apo and Holo Structures of Various Antibodies. Superposition of antibodies on their paratope. **(A)** aPCSK9-Fab2 displays minimal movement of CDR:H3 upon complexation of PCSK9, indicating a "lock and key" mechanism of binding. **(B)** Detailed view of the binding site of aDabi-Fab1.Both, apo and holo, conformations are seen during a MD simulation of the free Fab. As there are no major rearrangements, binding of dabigatran to aDabi-Fab1 is most likely achieved by "conformational selection". **(C)** aDabi-Fab2/3 and aPCSK9-Fab1 show comparable rearrangements of CDR:H3 upon binding to the antigen. In the apo structures the loop shields a hydrophobic binding pocket from solvent exposure.

7 Discussion - Conformational Flexibility of CDR:H3 in Various Antibodies

The antibodies presented in this work differ significantly in the conformational flexibility of their respective CDR:H3. It varies from no conformational changes upon binding for aPCSK9-Fab2, to minor rearrangements for aDabi-Fab1 but is most pronounced for aDabi-Fab2/3 and aPCSK9-Fab1. Regarding their conformational rearrangements seen in the crystal structures (see figure 40) and also taking into account the MD simulations, these antibodies represent models of a variety of binding mechanisms.

The binding of aPCSK9-Fab2 to PCSK9 may be regarded as a "lock and key" mechanism, first proposed by Emil Fischer in 1894 (Fischer, 1894). PCSK9 as well as the Fab do not display conformational changes upon binding compared to their apo forms. aDabi-Fab1 adopts different sidechain conformations for H:Trp52 in the apo and holo structures. H:Tyr100A and the apex of CDR:H3 and CDR:H2 also slightly change their conformations. MD simulations have shown that CDR:H3 as well as the two amino acids mentioned have a certain flexibility in the absence of dabigatran.

Both conformations are seen during MD, suggesting their presence in solution. It can thus be considered that dabigatran binds to aDabi-Fab by “conformational selection”, e.i. dabigatran selects the conformation preferred for binding from a variety of parallel existing low-energy conformational substates of the Fab (Kumar et al., 2000). aDabi-Fab2 and aPCSK9-Fab1 show the largest conformational rearrangements between their apo and holo structures, regarding CDR:H3. As explained above, it is unlikely that the respective conformations are due to crystal contacts. Interestingly, the conformational changes in CDR:H3 are quite similar in both Fabs, although they bind to very different antigens and also differ considerably in sequence. In both cases, a mostly hydrophobic cavity at the interface of heavy and light chain is shielded from contact with the polar water environment by the conformation of CDR:H3 adopted in the apo structures. CDR:H3 rearranges in both complex structures and the antigens bind into the cavity that becomes accessible. The binding partners then form H-bonds to a “hot spot” consisting of 2 - 3 amino acids in the otherwise hydrophobic cavity. These interactions have been shown to be critical for binding in both cases. MD simulations for aDabi-Fab2 and aPCSK9-Fab1 have shown that the CDR:H3 loops of the free Fabs have some intrinsic flexibility, that allows for access to the “hot spot”, but fail to adopt conformations observed in the complex structures. It is thus reasonable to assume a combination of “conformational selection” and “induced fit” (Koshland Jr, 1958) mechanisms: Besides the conformations seen in the crystal apo structures, there exist several additional low-energy substates, some of which allow for interaction between the antigen and the “hot spot” in the antibody’s paratope (“conformational selection”). After the formation of an initial encounter complex, binding induces rearrangements and additional interactions are formed. These finally result in the conformations seen in the respective complex crystal structures that would be energetically unfavored in the absence of antigen (“induced fit”). Although “conformational selection” was proposed to be the dominant mechanism for antibodies binding haptens (Okazaki and Takada, 2008), both mechanisms have been reported for a variety of antibodies (James et al., 2003; Estabrook and Reich, 2006; Rini et al., 1992; Wilson and Stanfield, 1994).

In antibodies, CDR:H3 is long known to be the CDR loop with the highest variability in sequence and conformation (Al-Lazikani et al., 1997). In fact, although some canonical structure classes have been proposed for CDR:H3 (Morea et al., 1998), a reliable classification is not possible to date. As presented in this work, its structural features allow for binding of antigens by several mechanisms, including “lock and key”, “conformational selection” and a combination of “conformational selection” and “induced fit”. It remains to be investigated, whether any of these mechanisms will prove beneficial for therapeutic antibodies against certain types of antigens.

References

M. Abifadel, M. Varret, J.P. Rabes, D. Allard, K. Ouguerram, M. Devillers, C. Cruaud, S. Benjannet, L. Wickham, and D. Erlich. Mutations in pcsk9 cause autosomal dominant hypercholesterolemia. *Nature genetics*, 34(2):154–156, 2003.

G.P. Adams, R. Schier, A.M. McCall, H.H. Simmons, E.M. Horak, R.K. Alpaugh, J.D. Marks, and L.M. Weiner. High affinity restricts the localization and tumor penetration of single-chain fv antibody molecules. *Cancer research*, 61(12):4750–4755, 2001.

A. Aderem and D.M. Underhill. Mechanisms of phagocytosis in macrophages. *Annual review of immunology*, 17(1):593–623, 1999.

B. Al-Lazikani, A.M. Lesk, and C. Chothia. Standard conformations for the canonical structures of immunoglobulins1. *Journal of Molecular Biology*, 273(4):927–948, 1997. ISSN 0022-2836.

B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J.D. Watson. Molecular biology of the cell. *Garland, New York*, 2002.

W.E. Alborn, G. Cao, H.E. Careskey, Y.W. Qian, D.R. Subramaniam, J. Davies, E.M. Conner, and R.J. Konrad. Serum proprotein convertase subtilisin kexin type 9 is correlated directly with serum ldl cholesterol. *Clinical chemistry*, 53(10):1814–1819, 2007.

D. Allard, S. Amsellem, M. Abifadel, M. Trillard, M. Devillers, G. Luc, M. Krempf, Y. Reznik, J.P. Girardet, and A. Fredenrich. Novel mutations of the pcsk9 gene cause variable phenotype of autosomal dominant hypercholesterolemia. *Human mutation*, 26(5):497–497, 2005. ISSN 1098-1004.

J. Ansell, J. Hirsh, L. Poller, H. Bussey, A. Jacobson, and E. Hylek. The pharmacology and management of the vitamin k antagonists: the seventh accp conference on antithrombotic and thrombolytic therapy. *Chest*, 126(3 Suppl):204S–, 2004. ISSN 0012-3692.

E.M. Antman, T.L. Wenger, V.P. Butler Jr, E. Haber, and T.W. Smith. Treatment of 150 cases of life-threatening digitalis intoxication with digoxin-specific fab antibody fragments. final report of a multicenter study. *Circulation*, 81(6):1744–1752, 1990.

T. Arakawa, J.S. Philo, K. Tsumoto, R. Yumioka, and D. Ejima. Elution of antibodies from a protein-a column by aqueous arginine solutions. *Protein expression and purification*, 36(2):244–248, 2004.

U.W. Arndt, J.N. Champness, R.P. Phizackerley, and A.J. Wonacott. A single-crystal oscillation camera for large unit cells. *Journal of Applied Crystallography*, 6(6): 457–463, 1973.

K. Bailey, F.R. Bettelheim, L. Lorand, and W.R. Middlebrook. Action of thrombin in the clotting of fibrinogen. 1951.

L. Baldi, D.L. Hacker, M. Adam, and F.M. Wurm. Recombinant protein production by large-scale transient gene expression in mammalian cells: state of the art and future perspectives. *Biotechnology letters*, 29(5):677–684, 2007. ISSN 0141-5492.

F.J. Baud, A. Sabouraud, E. Vicaut, P. Taboulet, J. Lang, C. Bismuth, J.M. Rouzioux, and J.M. Scherrmann. Treatment of severe colchicine overdose with colchicine-specific fab fragments. *New England Journal of Medicine*, 332(10):642–645, 1995.

C. Becattini, G. Agnelli, J. Emmerich, A. Bura, and J.I. Weitz. Initial treatment of venous thromboembolism. *Thrombosis and Haemostasis*, 96(3):242–, 2006. ISSN 0340-6245.

A. Beer. Bestimmung der absorption des rothen lichts in farbigen flüssigkeiten. *Ann. Phys. Chem*, 86(2):78–90, 1852.

S. Benjannet, D. Rhainds, R. Essalmani, J. Mayne, L. Wickham, W. Jin, M.C. Asselin, J. Hamelin, M. Varret, and D. Allard. Narc-1/pcsk9 and its natural mutants. *J.Biol.Chem.*, 279(47):48865–48875, 2004. ISSN 0021-9258.

H.C. Bimboim and J. Doly. A rapid alkaline extraction procedure for screening recombinant plasmid dna. *Nucleic acids research*, 7(6):1513–1523, 1979. ISSN 0305-1048.

R.C. Blake, A.R. Pavlov, and D.A. Blake. Automated kinetic exclusion assays to quantify protein binding interactions in homogeneous solution. *Analytical biochemistry*, 272(2):123–134, 1999.

S. Blech, T. Ebner, E. Ludwig-Schwellinger, J. Stangier, and W. Roth. The metabolism and disposition of the oral direct thrombin inhibitor, dabigatran, in humans. *Drug Metabolism and Disposition*, 36(2):386–, 2008. ISSN 0090-9556.

D. Blow. Outline of crystallography for biologists. 2003.

A.D. Boardman. *Electromagnetic surface modes*. Wiley, 1982.

W. Bode, I. Mayr, U. Baumann, R. Huber, R., S.R. Stone, and J. Hofsteenge. The refined 1.9 a crystal structure of human alpha-thrombin: interaction with d-phe-pro-arg chloromethylketone and significance of the tyr-pro-pro-trp insertion segment. *The EMBO journal*, 8(11):3467–, 1989.

F. Bollin, V. Dechavanne, and L. Chevalet. Design of experiment in cho and hek transient transfection condition optimization. *Protein Expression and Purification*, 2011. ISSN 1046-5928.

P. Bork, L. Holm, and C. Sander. The immunoglobulin fold. *J.Mol.Biol*, 242:309–320, 1994.

J.A. Bornhorst and J.J. Falke. [16] purification of proteins using polyhistidine affinity tags. *Methods in enzymology*, 326:245–254, 2000.

M.J. Bottomley, A. Cirillo, L. Orsatti, L. Ruggeri, T.S. Fisher, J.C. Santoro, R.T. Cummings, R.M. Cubbon, P.L. Surdo, and A. Calzetta. Structural and biochemical characterization of the wild type pcsk9-egf (ab) complex and natural familial hypercholesterolemia mutants. *Journal of Biological Chemistry*, 284(2):1313–1323, 2009. ISSN 0021-9258.

M. Brandl, M.S. Weiss, A. Jabs, J. Suehnel, and R. Hilgenfeld. Ch [pi]-interactions in proteins. *Journal of Molecular Biology*, 307(1):357–377, 2001. ISSN 0022-2836.

H. Brandstetter, D. Turk, H.W. Hoeffken, D. Grosse, J. Stuerzebecher, P.D. Martin, B.F.P. Edwards, and W. Bode. Refined 2.3 a x-ray crystal structure of bovine thrombin complexes formed with the benzamidine and arginine-based thrombin inhibitors napap, 4-tapap and mqpa: A starting point for improving antithrombotics. *Journal of Molecular Biology*, 226(4):1085–1099, 1992. ISSN 0022-2836.

S.E. Brenner, C. Chothia, and T.J.P. Hubbard. Population statistics of protein structures: lessons from structural classifications. *Current opinion in structural biology*, 7(3):369–376, 1997.

G. Bricogne. Molecular replacement. In *Proceedings of the CCP4 Study Weekend*, pages 62–75, 1992.

G. Bricogne and J. Irwin. Macromolecular refinement: Proceedings of the ccp4 study weekend, edited by e. *Dodson, M. Moore, A. Ralph & S. Bailey*, pages 85–92, 1996.

H.D. Brightbill, D.H. Libraty, S.R. Krutzik, R.B. Yang, J.T. Belisle, J.R. Bleharski, M. Maitland, M.V. Norgard, S.E. Plevy, S.T. Smale, et al. Host defense mecha-

nisms triggered by microbial lipoproteins through toll-like receptors. *Science*, 285 (5428):732–736, 1999.

M.S. Brown and J.L. Goldstein. A receptor-mediated pathway for cholesterol homeostasis. *Science*, 232(4746):34–47, 1986.

A.T. Bruenger. Free r value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, 355:472–475, 1992.

G.J. Brunn, D.E. Keyler, S.M. Pond, and P.R. Pentel. Reversal of desipramine toxicity in rats using drug-specific antibody fab'fragment: effects on hypotension and interaction with sodium bicarbonate. *Journal of Pharmacology and Experimental Therapeutics*, 260(3):1392–1399, 1992.

I.J. Bruno, J.C. Cole, M. Kessler, J. Luo, W.D.S. Motherwell, L.H. Purkis, B.R. Smith, R. Taylor, R.I. Cooper, S.E. Harris, et al. Retrieval of crystallographically-derived molecular geometry information. *Journal of chemical information and computer sciences*, 44(6):2133–2144, 2004.

S.K. Burley and G.A. Petsko. Aromatic-aromatic interaction: a mechanism of protein structure stabilization. *Science*, 229(4708):23–28, 1985. ISSN 0036-8075.

S.K. Burley and G.A. Petsko. Amino-aromatic interactions in proteins. *FEBS letters*, 203(2):139–143, 1986. ISSN 0014-5793.

G. Bussi, D. Donadio, and M. Parrinello. Canonical sampling through velocity-rescaling. *arXiv preprint arXiv:0803.4060*, 2008.

B. Cariou, K. Ouguerram, Y. Zaïr, R. Guerois, C. Langhi, S. Kourimate, I. Benoit, C. Le May, C. Gayet, K. Belabbas, et al. Pcsk9 dominant negative mutant results in increased ldl catabolic rate and familial hypobetalipoproteinemia. *Arteriosclerosis, thrombosis, and vascular biology*, 29(12):2191–2197, 2009.

J.C. Carrington and W.G. Dougherty. A viral cleavage site cassette: identification of amino acid sequences required for tobacco etch virus polyprotein processing. *Proceedings of the National Academy of Sciences*, 85(10):3391–3395, 1988.

P.J. Carter. Potent antibody therapeutics by design. *Nature Reviews Immunology*, 6 (5):343–357, 2006. ISSN 1474-1733.

C.W. Carter Jr and C.W. Carter. Protein crystallization using incomplete factorial experiments. *Journal of Biological Chemistry*, 254(23):12219–12223, 1979.

P. Chames, M. Van Regenmortel, E. Weiss, and D. Baty. Therapeutic antibodies: successes, limitations and hopes for the future. *British journal of pharmacology*, 157 (2):220–233, 2009.

J.C. Chan, D.E. Piper, Q. Cao, D. Liu, C. King, W. Wang, J. Tang, Q. Liu, J. Higbee, Z. Xia, Y. Di, S. Shetterly, Z. Arimura, H. Salomonis, W.G. Romanow, S.T. Thibault, R. Zhang, P. Cao, X.P. Yang, T. Yu, M. Lu, M.W. Retter, G. Kwon, K. Henne, O. Pan, M.M. Tsai, B. Fuchslocher, E. Yang, L. Zhou, K.J. Lee, M. Daris, J. Sheng, Y. Wang, W.D. Shen, W.C. Yeh, M. Emery, N.P. Walker, B. Shan, M. Schwarz, and S.M. Jackson. A proprotein convertase subtilisin/kexin type 9 neutralizing antibody reduces serum cholesterol in mice and nonhuman primates. *Proceedings of the National Academy of Sciences*, 106(24):9820–9825, June 2009. ISSN 1091-6490.

V.B. Chen, W.B. Arendall, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S. Richardson, and D.C. Richardson. Molprobity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography*, 66(1):12–21, 2009. ISSN 0907-4449.

A.J. Chu. Blood coagulation as an intrinsic pathway for proinflammation: a mini review. *Inflammation & Allergy-Drug Targets*, 9(1):32–44, 2010.

J.C. Cohen, E. Boerwinkle, T.H. Mosley Jr, and H.H. Hobbs. Sequence variations in pcsk9, low ldl, and protection against coronary heart disease. *New England Journal of Medicine*, 354(12):1264–1272, 2006. ISSN 0028-4793.

M.L. Connolly et al. Solvent-accessible surfaces of proteins and nucleic acids. *Science*, 221(4612):709–713, 1983.

S.J. Connolly, M.D. Ezekowitz, S. Yusuf, J. Eikelboom, J. Oldgren, A. Parekh, J. Pogue, P.A. Reilly, E. Themeles, and J. Varrone. Dabigatran versus warfarin in patients with atrial fibrillation. *New England Journal of Medicine*, 361(12):1139–1151, 2009. ISSN 0028-4793.

R. Cudney, S. Patel, K. Weisgraber, Y. Newhouse, and A. McPherson. Screening and optimization strategies for macromolecular crystal growth. *Acta Crystallographica Section D: Biological Crystallography*, 50(4):414–423, 1994.

D. Cunningham, D.E. Danley, K.F. Geoghegan, M.C. Griffor, J.L. Hawkins, T.A. Subashi, A.H. Varghese, M.J. Ammirati, J.S. Culp, and L.R. Hoth. Structural and biophysical studies of pcsk9 and its mutants linked to familial hypercholesterolemia. *Nature structural & molecular biology*, 14(5):413–419, 2007. ISSN 1545-9993.

T. Darkow, A.M. Vanderplas, K.H. Lew, J. Kim, and O. Hauch. Treatment patterns and real-world effectiveness of warfarin in nonvalvular atrial fibrillation within a managed care system. *Current Medical Research and Opinion*, 21(10):1583–1594, 2005. ISSN 0300-7995.

R.J. Darling and P.A. Brault. Kinetic exclusion assay technology: characterization of molecular interactions. *Assay and drug development technologies*, 2(6):647–657, 2004.

E.W. Davie. Biochemical and molecular aspects of the coagulation cascade. *Thrombosis and haemostasis*, 74(1):1, 1995.

C.G. Davis, J.L. Goldstein, T.C. Suedhof, R.G. Anderson, D.W. Russell, and M.S. Brown. Acid-dependent ligand dissociation and recycling of ldl receptor mediated by growth factor homology region. *Nature*, 326(6115):760–, 1987. ISSN 0028-0836.

G. De Crescenzo, C. Boucher, Y. Durocher, and M. Jolicoeur. Kinetic characterization by surface plasmon resonance-based biosensors: principle and emerging trends. *Cellular and Molecular Bioengineering*, 1(4):204–215, 2008. ISSN 1865-5025.

P. Debye. Interference of x-rays and thermal motion. *Annalen der Physik*, 43:49–95, 1914.

J. Deisenhofer. Crystallographic refinement and atomic models of a human fc fragment and its complex with fragment b of protein a from staphylococcus aureus at 2.9-and 2.8-. ang. resolution. *Biochemistry*, 20(9):2361–2370, 1981.

M. Di Nisio, S. Middeldorp, and H.R. Bueller. Direct thrombin inhibitors [review]. *New Engl J Med*, 353(10):1028–1040, 2005. ISSN 0028-4793.

K. Diederichs, P.A. Karplus, et al. Improved r-factors for diffraction data analysis in macromolecular crystallography. *Nature structural biology*, 4(4):269–275, 1997.

F. DiMaio, T.C. Terwilliger, R.J. Read, A. Wlodawer, G. Oberdorfer, U. Wagner, E. Valkov, A. Alon, D. Fass, H.L. Axelrod, et al. Improved molecular replacement by density-and energy-guided protein structure optimization. *Nature*, 473(7348): 540–543, 2011.

B. Dong, M. Wu, H. Li, F.B. Kraemer, K. Adeli, N.G. Seidah, S.W. Park, and J. Liu. Strong induction of pcsk9 gene expression through hnf1alpha and srebp2: mechanism for the resistance to ldl-cholesterol lowering effect of statins in dyslipidemic hamsters. *Journal of lipid research*, 51(6):1486–1495, 2010. ISSN 0022-2275.

W.G. Dougherty, J.C. Carrington, S.M. Cary, and T.D. Parks. Biochemical and mutational analysis of a plant virus polyprotein cleavage site. *The EMBO journal*, 7(5): 1281, 1988.

G. Dubuc, M. Tremblay, G. Paré, H. Jacques, J. Hamelin, S. Benjannet, L. Boulet, J. Genest, L. Bernier, N.G. Seidah, et al. A new method for measurement of total plasma pcsk9: clinical applications. *Journal of lipid research*, 51(1):140–149, 2010.

C.J. Duff, M.J. Scott, I.T. Kirby, S.E. Hutchinson, S.L. Martin, and N.M. Hooper. Antibody-mediated disruption of the interaction between pcsk9 and the low-density lipoprotein receptor. *Biochemical Journal*, 419(Pt 3):577–, 2009.

Y. Durocher, S. Perret, and A. Kamen. High-level and high-throughput recombinant protein production by transient transfection of suspension-growing human 293-ebna1 cells. *Nucleic acids research*, 30(2):e9–e9, 2002. ISSN 0305-1048.

H. Edelhoch. Spectroscopic determination of tryptophan and tyrosine in proteins*. *Biochemistry*, 6(7):1948–1954, 1967.

G.M. Edelman and M.D. Poulik. Studies on structural units of the γ-globulins. *The Journal of experimental medicine*, 113(5):861–884, 1961.

P. Emsley and K. Cowtan. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D: Biological Crystallography*, 60(12):2126–2132, 2004. ISSN 0907-4449.

R.A. Engh and R. Huber. Accurate bond and angle parameters for x-ray protein structure refinement. *Acta Crystallographica Section A: Foundations of Crystallography*, 47(4):392–400, 1991.

E. Engvall, P. Perlmann, et al. Enzyme-linked immunosorbent assay (elisa). quantitative assay of immunoglobulin g. *Immunochemistry*, 8(9):871, 1971.

D.A. Erlanson, R.S. McDowell, T. O'Brien, et al. Fragment-based drug discovery. *Journal of medicinal chemistry*, 47(14):3463–3482, 2004.

U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, and L.G. Pedersen. A smooth particle mesh ewald method. *The Journal of Chemical Physics*, 103:8577, 1995.

R.A. Estabrook and N. Reich. Observing an induced-fit mechanism during sequence-specific dna methylation. *Journal of Biological Chemistry*, 281(48):37205–37214, 2006. ISSN 0021-9258.

P. Evans. Scala. *Medical Research Council, CCP 4 Newsletter on Protein Crystallography*, pages 22–24, 1997.

P. Evans. Scaling and assessment of data quality. *Acta Crystallographica Section D: Biological Crystallography*, 62(1):72–82, 2005.

P.P. Ewald. Das reziproke gitter in der strukturtheorie. *Zeitschrift fuer Kristallographie*, 56:129–156, 1921.

P.L. Felgner, T.R. Gadek, M. Holm, R. Roman, H.W. Chan, M. Wenz, J.P. Northrop, G.M. Ringold, and M. Danielsen. Lipofection: a highly efficient, lipid-mediated dna-transfection procedure. *Proceedings of the National Academy of Sciences*, 84 (21):7413–, 1987. ISSN 0027-8424.

M.J. Field, M. Albe, C. Bret, P.D. Martin, A. Thomas, et al. The dynamo library for molecular simulations using hybrid quantum mechanical and molecular mechanical potentials. *Journal of Computational Chemistry*, 21(12):1088–1100, 2000.

L.D. Fiore, M.A. Scola, C.E. Cantillon, and M.T. Brophy. Anaphylactoid reactions to vitamin k. *Journal of thrombosis and thrombolysis*, 11(2):175–183, 2001. ISSN 0929-5305.

E. Fischer. Einfluss der configuration auf die wirkung der enzyme.(1894) ber. *Berichte der deutschen chemischem Gesellschaft*, 27:2984, 1894.

T.S. Fisher, P.L. Surdo, S. Pandit, M. Mattu, J.C. Santoro, D. Wisniewski, R.T. Cummings, A. Calzetta, R.M. Cubbon, P.A. Fischer, et al. Effects of ph and low density lipoprotein (ldl) on pcsk9-dependent ldl receptor regulation. *Journal of Biological Chemistry*, 282(28):20502–20512, 2007.

A. Forsgren and J. Sjöquist. "Protein A" from S. Aureus. *The Journal of Immunology*, 97(6):822–827, 1966.

M. Frank-Kamenetsky, A. Grefhorst, N.N. Anderson, T.S. Racie, B. Bramlage, A. Akinc, D. Butler, K. Charisse, R. Dorkin, Y. Fan, et al. Therapeutic rnai targeting pcsk9 acutely lowers plasma cholesterol in rodents and ldl cholesterol in nonhuman primates. *Proceedings of the National Academy of Sciences*, 105(33):11915–11920, 2008.

M.W. Freyer and E.A. Lewis. Isothermal titration calorimetry: experimental design, data analysis, and probing macromolecule/ligand binding and kinetic interactions. *Methods in cell biology*, 84:79–113, 2008.

D. Gailani, G.J. Broze Jr, et al. Factor xi activation in a revised model of blood coagulation. *Science (New York, NY)*, 253(5022):909, 1991.

D.J. Galbraith, A.S. Tait, A.J. Racher, J.R. Birch, and D.C. James. Control of culture environment for improved polyethylenimine-mediated transient production of recombinant monoclonal antibodies by cho cells. *Biotechnology progress*, 22(3): 753–762, 2006.

T. Gallagher, G. Gilliland, L. Wang, and P. Bryan. The prosegment-subtilisin bpn' complex: crystal structure of a specific 'foldase'. *Structure*, 3(9):907–914, 1995.

J.P. Gallivan and D.A. Dougherty. Cation-pi interactions in structural biology. *Proceedings of the National Academy of Sciences*, 96(17):9459–, 1999. ISSN 0027-8424.

E.F. Garman and R.L. Owen. Cryocooling and radiation damage in macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography*, 62(1): 32–47, 2005.

P. Gasque et al. Complement: a unique innate immune sensor for danger signals. *Molecular immunology*, 41(11):1089, 2004.

R. Geisberger, M. Lamers, and G. Achatz. The riddle of the dual expression of igm and igd. *Immunology*, 118(4):429–437, 2006.

S.C. Gill, P.H. von Hippel, et al. Calculation of protein extinction coefficients from amino acid sequence data. *Analytical biochemistry*, 182(2):319–326, 1989.

J.W. Goding, D.W. Scott, and J.E. Layton. Genetics, cellular expression and function of igd and igm receptors. *Immunological reviews*, 37(1):152–180, 1977.

S. Gordon. The role of the macrophage in immune regulation. *Research in immunology*, 149(7):685–688, 1998.

Y. Goto, T. Azuma, and K. Hamaguchi. Refolding of the immunoglobulin light chain. *Journal of biochemistry*, 85(6):1427–1438, 1979.

J. Goudswaard, JA Donk, A. Noordzij, R.H. Dam, and J.P. Vaerman. Protein a reactivity of various mammalian immunoglobulins. *Scandinavian journal of immunology*, 8(1):21–28, 2006.

D.W. Green, V.M. Ingram, and M.F. Perutz. The structure of haemoglobin. iv. sign determination by the isomorphous replacement method. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 225(1162):287–307, 1954.

L.L. Green, M.C. Hardy, C.E. Maynard-Currie, H. Tsuda, D.M. Louie, M.J. Mendez, H. Abderrahim, M. Noguchi, D.H. Smith, Y. Zeng, et al. Antigen–specific human monoclonal antibodies from mice engineered with human ig heavy and light chain yacs. *Nature genetics*, 7(1):13–21, 1994.

U. Grenander. *Probability and statistics: the Harald Cramér volume*. Almqvist & Wiksell, 1959.

N. Gupta, N. Fisker, M.C. Asselin, M. Lindholm, C. Rosenbohm, H. Ørum, J. Elmén, N.G. Seidah, and E.M. Straarup. A locked nucleic acid antisense oligonucleotide (lna) silences pcsk9 and enhances ldlr expression in vitro and in vivo. *PloS one*, 5 (5):e10682, 2010.

T. Hahn, U. Shmueli, A.J.C. Wilson, and E. Prince. *International tables for crystallography*. D. Reidel Publishing Company, 2005.

R. Haldankar, D. Li, Z. Saremi, C. Baikalov, and R. Deshpande. Serum-free suspensin large-scale transient transfection of cho cells in wave bioreactors. *Molecular Biotechnology*, 34(2):191–199, 2006. ISSN 1073-6085.

E.N. Hampton, M.W. Knuth, J. Li, J.L. Harris, S.A. Lesley, and G. Spraggon. The self-inhibited structure of full-length pcsk9 at 1.9 a reveals structural homology with resistin within the c-terminal domain. *Proceedings of the National Academy of Sciences*, 104(37):14604–, 2007. ISSN 0027-8424.

S. Hanessian, G. Papeo, K. Fettis, E. Therrien, and M.T.P. Viet. Synthesis of 310-helix-inducing constrained analogues of l-proline. *The Journal of Organic Chemistry*, 69 (15):4891–4899, 2004. ISSN 0022-3263.

L.J. Harris, E. Skaletsky, and A. McPherson. Crystallization of intact monoclonal antibodies. *Proteins: Structure, Function, and Bioinformatics*, 23(2):285–289, 2004.

N.H. Hauel, H. Nar, H. Priepke, U. Ries, J.M. Stassen, and W. Wienen. Structure-based design of novel potent nonpeptide thrombin inhibitors. *Journal of medicinal chemistry*, 45(9):1757–1766, 2002. ISSN 0022-2623.

B. Henrich, A. Bergamaschi, C. Broennimann, R. Dinapoli, E.F. Eikenberry, I. Johnson, M. Kobas, P. Kraft, A. Mozzanica, and B. Schmitt. Pilatus: A single photon counting pixel detector for x-ray applications. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 607(1):247–249, 2009.

B. Hess. P-lincs: A parallel linear constraint solver for molecular simulation. *Journal of Chemical Theory and Computation*, 4(1):116–122, 2008.

B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl. Gromacs 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of chemical theory and computation*, 4(3):435–447, 2008.

J. Hirsh, J.E. Dalen, D.R. Anderson, L. Poller, H. Bussey, J. Ansell, and D. Deykin. Oral anticoagulants: mechanism of action, clinical effectiveness, and optimal therapeutic range. *Chest*, 119(1 suppl):8S–21S, 2001a. ISSN 0012-3692.

J. Hirsh, T.E. Warkentin, S.G. Shaughnessy, S.S. Anand, J.L. Halperin, R. Raschke, C. Granger, E.M. Ohman, and J.E. Dalen. Heparin and low-molecular-weight heparin mechanisms of action, pharmacokinetics, dosing, monitoring, efficacy, and safety. *Chest*, 119(1 suppl):64S–94S, 2001b. ISSN 0012-3692.

E. Hochuli, H. Döbeli, and A. Schacher. New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues. *Journal of Chromatography A*, 411:177–184, 1987.

A. Hoffmann-Röder, E. Schweizer, J. Egger, P. Seiler, U. Obst-Sander, B. Wagner, M. Kansy, D.W. Banner, and F. Diederich. Mapping the fluorophilicity of a hydrophobic pocket: Synthesis and biological evaluation of tricyclic thrombin inhibitors directing fluorinated alkyl groups into the p pocket. *ChemMedChem*, 1 (11):1205–1215, 2006.

H. Hope. Introduction to cryocrystallography. *International Tables for Crystallography*, 2006.

J.D. Horton, J.C. Cohen, and H.H. Hobbs. Molecular biology of pcsk9: its role in ldl metabolism. *Trends in biochemical sciences*, 32(2):71–77, 2007. ISSN 0968-0004.

G.C. Ireton and B.L. Stoddard. Microseed matrix screening to improve crystals of yeast cytosine deaminase. *Acta Crystallographica Section D: Biological Crystallography*, 60(3):601–605, 2004. ISSN 0907-4449.

I.I. Ivanov, R.L. Schelonka, Y. Zhuang, G.L. Gartland, M. Zemlin, and H.W. Schroeder Jr. Development of the expressed ig cdr-h3 repertoire is marked by focusing of constraints in length, amino acid use, and charge that are first established in early b cell progenitors. *The Journal of Immunology*, 174(12):7773–7780, 2005.

L.C. James, P. Roversi, and D.S. Tawfik. Antibody multispecificity mediated by conformational diversity. *Science*, 299(5611):1362–1367, 2003. ISSN 0036-8075.

J. Jancarik and S.H. Kim. Sparse matrix sampling: a screening method for crystallization of proteins. *Journal of applied crystallography*, 24(4):409–411, 1991.

C.A. Janeway, P. Travers, M. Walport, and J.D. Capra. *Immunobiology: the immune system in health and disease*, volume 1. Current Biology, 2001.

G.A. Jeffrey and W. Saenger. *Hydrogen bonding in biological structures*. Springer-Verlag Berlin, 1994.

H. Jeon and S.C. Blacklow. Structure and physiologic function of the low-density lipoprotein receptor. *Annu.Rev.Biochem.*, 74:535–562, 2005. ISSN 0066-4154.

G. Johnson and T. Te Wu. Kabat database and its applications: future directions. *Nucleic acids research*, 29(1):205–206, 2001. ISSN 0305-1048.

B. Johnsson, S. Löfås, and G. Lindquist. Immobilization of proteins to a carboxymethyldextran-modified gold surface for biospecific interaction analysis in surface plasmon resonance sensors. *Analytical biochemistry*, 198(2):268–277, 1991.

M. Jones, P. McEwan, C.L. Morgan, J.R. Peters, J. Goodfellow, and C.J. Currie. Evaluation of the pattern of treatment, level of anticoagulation control, and outcome of treatment with warfarin in patients with non-valvar atrial fibrillation: a record linkage study in a large british population. *Heart*, 91(4):472–477, 2005. ISSN 1468-201X.

P.T. Jones, P.H. Dear, J. Foote, M.S. Neuberger, and G. Winter. Replacing the complementarity-determining regions in a human antibody with those from a mouse. *Nature*, 321:522 – 525, 1986.

W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, and M.L. Klein. Comparison of simple potential functions for simulating liquid water. *The Journal of chemical physics*, 79:926, 1983.

P.D. Josephy, T. Eling, and R.P. Mason. The horseradish peroxidase-catalyzed oxidation of 3, 5, 3', 5'-tetramethylbenzidine. free radical and charge-transfer complex intermediates. *Journal of Biological Chemistry*, 257(7):3669–3675, 1982.

W. Kabsch. Xds. *Acta Crystallographica Section D: Biological Crystallography*, 66(2): 125–132, 2010.

W. Kabsch and C. Sander. Dssp: definition of secondary structure of proteins given a set of 3d coordinates. *Biopolymers*, 22:2577–2637, 1983.

K.A. Kantardjieff and B. Rupp. Matthews coefficient probabilities: improved estimates for unit cell contents of proteins, dna, and protein–nucleic acid complex crystals. *Protein Science*, 12(9):1865–1871, 2009.

R. Karlsson and A. Fält. Experimental design for kinetic analysis of protein-protein interactions with surface plasmon resonance biosensors. *Journal of immunological methods*, 200(1):121–133, 1997.

C.W. Khoo, K.H. Tay, E. Shantsila, and G.Y.H. Lip. Novel oral anticoagulants. *International journal of clinical practice*, 63(4):630–641, 2009. ISSN 1742-1241.

G.J. Kleywegt, A.T. Brünger, et al. Checking your imagination: applications of the free r value. *Structure*, 4(8):897–904, 1996.

H. Kobayashi, H. Sakahara, K. Endo, M. Hosono, Z.S. Yao, S. Toyama, and J. Konishi. Comparison of the chase effects of avidin, streptavidin, neutravidin, and avidin-ferritin on a radiolabeled biotinylated anti-tumor monoclonal antibody. *Cancer Science*, 86(3):310–314, 1995.

G. Köhler and C. Milstein. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature*, 256(5517):495–497, 1975.

I. Kola and J. Landis. Can the pharmaceutical industry reduce attrition rates? *Nature reviews Drug discovery*, 3(8):711–716, 2004.

D.E. Koshland Jr. Application of a theory of enzyme specificity to protein synthesis. *Proceedings of the National Academy of Sciences of the United States of America*, 44 (2):98–, 1958.

I.K. Kotowski, A. Pertsemlidis, A. Luke, R.S. Cooper, G.L. Vega, J.C. Cohen, and H.H. Hobbs. A spectrum of pcsk9 alleles contributes to plasma levels of low-density lipoprotein cholesterol. *The American Journal of Human Genetics*, 78(3):410–422, 2006. ISSN 0002-9297.

E. Kretschmann and H. Raether. Radiative decay of non radiative surface plasmons excited by light(surface plasma waves excitation by light and decay into photons applied to nonradiative modes). *Zeitschrift Fuer Naturforschung, Teil A*, 23:2135–, 1968.

E. Krissinel and K. Henrick. Secondary-structure matching (ssm), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallographica Section D: Biological Crystallography*, 60(12):2256–2268, 2004. ISSN 0907-4449.

E. Krissinel and K. Henrick. Inference of macromolecular assemblies from crystalline state. *Journal of Molecular Biology*, 372(3):774–797, 2007. ISSN 0022-2836.

S. Kumar, B. Ma, C.J. Tsai, N. Sinha, and R. Nussinov. Folding and binding cascades: dynamic landscapes and population shifts. *Protein Science*, 9(1):10–19, 2000. ISSN 1469-896X.

B.T. Kurien and R.H. Scofield. Common artifacts and mistakes made in electrophoresis. *Methods in molecular biology (Clifton, NJ)*, 869:633–, 2012. ISSN 1940-6029.

A. Kuusinen, M. Arvola, C. Oker-Blom, and K. Keinänen. Purification of recombinant glur-d glutamate receptor produced in sf21 insect cells. *European Journal of Biochemistry*, 233(3):720–726, 1995.

H.J. Kwon, T.A. Lagace, M.C. McNutt, J.D. Horton, and J. Deisenhofer. Molecular basis for ldl receptor recognition by pcsk9. *Proceedings of the National Academy of Sciences*, 105(6):1820–1825, February 2008. ISSN 1091-6490.

J.E. Ladbury and M.L. Doyle. Biocalorimetry 2. *Applications Calorimetry In The Biological Sciences. The Sussex John Wiley & Sons, Ltd*, page 259, 2004.

U.K. Laemmli et al. Cleavage of structural proteins during the assembly of the head of bacteriophage t4. *Nature*, 227(5259):680–685, 1970.

T.A. Lagace, D.E. Curtis, R. Garuti, M.C. McNutt, S.W. Park, H.B. Prather, N.N. Anderson, Y.K. Ho, R.E. Hammer, and J.D. Horton. Secreted pcsk9 decreases the number of ldl receptors in hepatocytes and inlivers of parabiotic mice. *Journal of Clinical Investigation*, 116(11):2995–3005, 2006.

K. Landsteiner. *The specificity of serological reactions*. Dover Publications, 1990.

G.H. Lathe and C.R. Ruthven. The separation of substances on the basis of their molecular weights, using columns of starch and water. *The Biochemical journal*, 60 (4):xxxiv, 1955.

A.G.W. Leslie. The integration of macromolecular diffraction data. *Acta Crystallographica Section D: Biological Crystallography*, 62(1):48–57, 2005.

C.E. Leysath, A.F. Monzingo, J.A. Maynard, J. Barnett, G. Georgiou, B.L. Iverson, and J.D. Robertus. Crystal structure of the engineered neutralizing antibody m18 complexed to domain 4 of the anthrax protective antigen. *Journal of Molecular Biology*, 387(3):680–693, 2009. ISSN 0022-2836.

Y. Li, H. Li, F. Yang, S.J. Smith-Gill, and R.A. Mariuzza. X-ray snapshots of the maturation of an antibody response to a protein antigen. *Nature structural biology*, 10 (6):482–488, 2003. ISSN 1072-8368.

H. Liang, J. Chaparro-Riggers, P. Strop, T. Geng, J.E. Sutton, D. Tsai, L. Bai, Y. Abdiche, J. Dilley, J. Yu, S. Wu, S.M. Chin, N.A. Lee, A. Rossi, J.C. Lin, A. Rajpal, J. Pons, and D.L. Shelton. Proprotein convertase substilisin/kexin type 9 antagonism reduces low-density lipoprotein cholesterol in statin-treated hypercholesterolemic nonhuman primates. *The Journal of pharmacology and experimental therapeutics*, 340(2):228–236, 2012. ISSN 00223565 (ISSN).

K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J.L. Klepeis, R.O. Dror, and D.E. Shaw. Improved side-chain torsion potentials for the amber ff99sb protein force field. *Proteins: Structure, Function, and Bioinformatics*, 78(8):1950–1958, 2010.

S. Löfås and B. Johnsson. A novel hydrogel matrix on gold surfaces in surface plasmon resonance sensors for fast and efficient covalent immobilization of ligands. *Journal of the Chemical Society, Chemical Communications*, (21):1526–1528, 1990.

N. Lonberg. Human antibodies from transgenic animals. *Nature biotechnology*, 23 (9):1117–1125, 2005.

F. Lottspeich and H. Zorbas. Bioanalytik. *Spektrum Akademischer Verlag*, 1998.

S.C. Lovell, J.M. Word, J.S. Richardson, and D.C. Richardson. The penultimate rotamer library. *Proteins: Structure, Function, and Bioinformatics*, 40(3):389–408, 2000.

D. Luo and W.M. Saltzman. Synthetic dna delivery systems. *Nature biotechnology*, 18 (1):33–37, 2000. ISSN 1087-0156.

R.M. MacCallum, A.C.R. Martin, and J.M. Thornton. Antibody-antigen interactions: contact analysis and binding site topography. *Journal of Molecular Biology*, 262(5): 732–745, 1996. ISSN 0022-2836.

M.G. Mage. [6] preparation of fab fragments from iggs of different animal species. *Methods in enzymology*, 70:142–150, 1980.

K.G. Mann and L. Lorand. Introduction: blood coagulation. *Methods in enzymology*, 222:1–10, 1993.

KG Mann, C. Van't Veer, K. Cawthern, S. Butenas, et al. The role of the tissue factor pathway in initiation of coagulation. *Blood coagulation & fibrinolysis: an international journal in haemostasis and thrombosis*, 9:S3, 1998.

C.D. Mathers and D. Loncar. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS medicine*, 3(11):e442–, 2006. ISSN 1549-1676.

B.W. Matthews. Solvent content of protein crystals. *J. Mol. Biol.*, 33:491–497, 1968.

K.N. Maxwell and J.L. Breslow. Adenoviral-mediated expression of pcsk9 in mice results in a low-density lipoprotein receptor knockout phenotype. *Proceedings of the National Academy of Sciences of the United States of America*, 101(18):7100–, 2004. ISSN 0027-8424.

G. Mayer, S. Poirier, and N.G. Seidah. Annexin a2 is a c-terminal pcsk9-binding protein that regulates endogenous low density lipoprotein receptor levels. *Journal of Biological Chemistry*, 283(46):31791–31801, 2008. ISSN 0021-9258.

J. McCafferty, A.D. Griffiths, G. Winter, and D.J. Chiswell. Phage antibodies: filamentous phage displaying antibody variable domains. *Nature*, 348, 1990.

A.J. McCoy. Liking likelihood. *Acta Crystallographica Section D: Biological Crystallography*, 60(12):2169–2183, 2004.

A.J. McCoy. Solving structures of protein complexes by molecular replacement with phaser. *Acta Crystallographica Section D: Biological Crystallography*, 63(1):32–41, 2006.

A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, and R.J. Read. Phaser crystallographic software. *Journal of applied crystallography*, 40(4): 658–674, 2007. ISSN 0021-8898.

G.B. McGaughey, M. Gagne, and A.K. Rappe. pi-stacking interactions. *Journal of Biological Chemistry*, 273(25):15458–15463, 1998. ISSN 0021-9258.

A. McPherson. *Crystallization of biological macromolecules*, volume 586. Cold Spring Harbor Laboratory Press New York, 1999.

P. Meissner, H. Pick, A. Kulangara, P. Chatellard, K. Friedrich, and F.M. Wurm. Transient gene expression: Recombinant protein production with suspension-adapted hek293-ebna cells. *Biotechnology and bioengineering*, 75(2):197–203, 2001.

S. Miyamoto and P.A. Kollman. Settle: an analytical version of the shake and rattle algorithm for rigid water models. *Journal of computational chemistry*, 13(8):952–962, 1992.

V. Morea, A. Tramontano, M. Rustici, C. Chothia, and A.M. Lesk. Conformations of the third hypervariable region in the vh domain of immunoglobulins1. *Journal of Molecular Biology*, 275(2):269–294, 1998. ISSN 0022-2836.

S. Muyldermans, T. Atarhouch, J. Saldanha, J. Barbosa, and R. Hamers. Sequence and structure of vh domain from naturally occurring camel heavy chain immunoglobulins lacking light chains. *Protein engineering*, 7(9):1129–1135, 1994.

D.G. Myszka, X. He, M. Dembo, T.A. Morton, and B. Goldstein. Extending the range of rate constants available from biacore: interpreting mass transport-influenced binding data. *Biophysical journal*, 75(2):583–594, 1998.

H. Nar. The role of structural information in the discovery of direct thrombin and factor xa inhibitors. *Trends in Pharmacological Sciences*, 2012. ISSN 0165-6147.

N. Nassoury, D.A. Blasiole, A. Tebon Oler, S. Benjannet, J. Hamelin, V. Poupon, P.S. McPherson, A.D. Attie, A. Prat, and N.G. Seidah. The cellular trafficking of the secretory proprotein convertase pcsk9 and its dependence on the ldlr. *Traffic*, 8(6): 718–732, 2007. ISSN 1600-0854.

S. Naureckiene, L. Ma, K. Sreekumar, U. Purandare, C. Frederick Lo, Y. Huang, L.W. Chiang, J.M. Grenier, B.A. Ozenberger, and J. Steven Jacobsen. Functional characterization of narc 1, a novel proteinase related to proteinase k. *Archives of biochemistry and biophysics*, 420(1):55–67, 2003. ISSN 0003-9861.

A.L. Nelson, E. Dhimolea, and J.M. Reichert. Development trends for human monoclonal antibody therapeutics. *Nature Reviews Drug Discovery*, 9(10):767–774, 2010.

MS Neuberger, GT Williams, EB Mitchell, SS Jouhal, JG Flanagan, and TH Rabbitts. A hapten-specific chimaeric ige antibody with human physiological effector function. *Nature*, 314(6008):268–270, 1985.

Y.G. Ni, J.H. Condra, L. Orsatti, X. Shen, S. Di Marco, S. Pandit, M.J. Bottomley, L. Ruggeri, R.T. Cummings, and R.M. Cubbon. A proprotein convertase subtilisin-like/kexin type 9 (pcsk9) c-terminal domain antibody antigen-binding fragment inhibits pcsk9 internalization and restores low density lipoprotein uptake. *Journal of Biological Chemistry*, 285(17):12882–, 2010. ISSN 0021-9258.

M.H. Niemi, L. Turunen, T. Pulli, T.K. Nevanen, M. Hoeyhtyae, H. Soederlund, J. Rouvinen, and K. Takkinen. A structural insight into the molecular recognition of a (-)-delta9-tetrahydrocannabinol and the development of a sensitive, one-step, homogeneous immunocomplex-based assay for its detection. *Journal of Molecular Biology*, 400(4):803–814, 2010. ISSN 0022-2836.

M.J. Niles, L. Matsuuchi, and M.E. Koshland. Polymer igm assembly and secretion in lymphoid and nonlymphoid cell lines: evidence that j chain is required for pentamer igm synthesis. *Proceedings of the National Academy of Sciences*, 92(7):2884–2888, 1995.

A. Nisonoff et al. Enzymatic digestion of rabbit gamma globulin and antibody and chromatography of digestion products. *Methods in medical research*, 10:134, 1964.

W. Niu, Z. Chen, P.S. Gandhi, A.D. Vogt, N. Pozzi, L.A. Pelc, F. Zapata, and E. Di Cera. Crystallographic and kinetic evidence of allostery in a trypsin-like protease. *Biochemistry*, 50(29):6301–6307, 2011.

A. Nurisso, A. Daina, and R.C. Walker. A practical introduction to molecular dynamics simulations: applications to homology modeling. *Methods in molecular biology (Clifton, NJ)*, 857:137, 2012.

R.J. Ober, C. Martinez, X. Lai, J. Zhou, and E.S. Ward. Exocytosis of igg as mediated by the receptor, fcrn: an analysis at the single-molecule level. *Proceedings of the National Academy of Sciences of the United States of America*, 101(30):11076–11081, 2004.

K. Okazaki and S. Takada. Dynamic energy landscape view of coupled binding and protein conformational change: Induced-fit versus population-shift mechanisms. *Proceedings of the National Academy of Sciences*, 105(32):11182–, 2008. ISSN 0027-8424.

A. Onell and K. Andersson. Kinetic determinations of molecular interactions using biacore–minimum data requirements for efficient experimental design. *Journal of Molecular Recognition*, 18(4):307–317, July 2005. ISSN 0952-3499.

C.N. Pace, F. Vajdos, L. Fee, G. Grimsley, and T. Gray. How to measure and predict the molar absorption coefficient of a protein. *Protein Science*, 4(11):2411–2423, 2008.

S.W. Park, Y.A. Moon, and J.D. Horton. Post-transcriptional regulation of low density lipoprotein receptor protein by proprotein convertase subtilisin/kexin type 9a in

mouse liver. *Journal of Biological Chemistry*, 279(48):50630–50638, 2004. ISSN 0021-9258.

M. Parrinello and A. Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied physics*, 52(12):7182–7190, 1981.

A.L. Patterson. A fourier series method for the determination of the components of interatomic distances in crystals. *Physical Review*, 46(5):372, 1934.

M.C. Peitsch and J. Tschopp. Assembly of macromolecular pores by immune defense systems. *Current opinion in cell biology*, 3(4):710–716, 1991.

E.C. Peterson, E.M. Laurenzana, W.T. Atchley, H.P. Hendrickson, and S.M. Owens. Development and preclinical testing of a high-affinity single-chain antibody against (+)-methamphetamine. *Journal of Pharmacology and Experimental Therapeutics*, 325(1):124–133, 2008.

PL Pham, S. Perret, B. Cass, E. Carpentier, G. St-Laurent, L. Bisson, A. Kamen, and Y. Durocher. Transient gene expression in hek293 cells: peptone addition posttransfection improves recombinant protein synthesis. *Biotechnology and bioengineering*, 90(3):332–344, 2005.

M.M. Pierce, CS Raman, B.T. Nall, et al. Isothermal titration calorimetry of protein-protein interactions. *Methods*, 19(2):213–221, 1999.

D.E. Piper, S. Jackson, Q. Liu, W.G. Romanow, S. Shetterly, S.T. Thibault, B. Shan, and N.P.C. Walker. The crystal structure of pcsk9: a regulator of plasma ldl-cholesterol. *Structure*, 15(5):545–552, 2007. ISSN 0969-2126.

L. Pisciotta, C.P. Oliva, A.B. Cefalu, D. Noto, A. Bellocchio, R. Fresa, A. Cantafora, D. Patel, M. Averna, and P. Tarugi. Additive effect of mutations in ldlr and pcsk9 genes on the phenotype of familial hypercholesterolemia. *Atherosclerosis*, 186(2): 433–440, 2006. ISSN 0021-9150.

S. Poirier, G. Mayer, V. Poupon, P.S. McPherson, R. Desjardins, K. Ly, M.C. Asselin, R. Day, F.J. Duclos, and M. Witmer. Dissection of the endogenous cellular pathways of pcsk9-induced low density lipoprotein receptor degradation. *Journal of Biological Chemistry*, 284(42):28856–28864, 2009. ISSN 0021-9258.

J. Porath, P. Flodin, et al. Gel filtration: a method for desalting and group separation. *Nature*, 183(4676):1657–1659, 1959.

J. Porath, J.A.N. Carlsson, I. Olsson, and G. Belfrage. Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature*, 258:598–599, 1975.

R. Porsche and Z.R. Brenner. Allergy to protamine sulfate. *Heart & Lung: The Journal of Acute and Critical Care*, 28(6):418–428, 1999. ISSN 0147-9563.

R.R. Porter. The formation of a specific inhibitor by hydrolysis of rabbit antiovalbumin. *Biochemical Journal*, 46(4):479, 1950.

R.R. Porter. The hydrolysis of rabbit γ-globulin and antibodies with crystalline papain. *Biochemical Journal*, 73(1):119, 1959.

R.R. Porter. Chemical structure of γ-globulin and antibodies. *British Medical Bulletin*, 19(3):197–201, 1963.

E. Pozharski, M.A. Wilson, A. Hewagama, A.B. Shanafelt, G. Petsko, and D. Ringe. Anchoring a cationic ligand: the structure of the fab fragment of the anti-morphine antibody 9b1 and its complex with morphine. *Journal of Molecular Biology*, 337 (3):691–697, 2004. ISSN 0022-2836.

F.W. Putnam, Y.S. Liu, and T.L. Low. Primary structure of a human iga1 immunoglobulin. iv. streptococcal iga1 protease, digestion, fab and fc fragments, and the complete amino acid sequence of the alpha 1 heavy chain. *Journal of Biological Chemistry*, 254(8):2865–2874, 1979.

H. Raether. Surface plasmons on smooth and rough surfaces and on gratings. *Springer Tracts in Modern Physics (Springer, Berlin, 1988)*, pages –, 1997.

G.N. Ramachandran, C. Ramakrishnan, and V. Sasisekharan. Stereochemistry of polypeptide chain configurations. *Journal of Molecular Biology*, 7:95, 1963.

S. Rashid, D.E. Curtis, R. Garuti, N.N. Anderson, Y. Bashmakov, YK Ho, R.E. Hammer, Y.A. Moon, and J.D. Horton. Decreased plasma cholesterol and hypersensitivity to statins in mice lacking pcsk9. *Proceedings of the National Academy of Sciences of the United States of America*, 102(15):5374–5379, 2005.

P. Rathanaswami, S. Roalstad, L. Roskos, Q.J. Su, S. Lackie, and J. Babcook. Demonstration of an in vivo generated sub-picomolar affinity fully human monoclonal antibody to interleukin-8. *Biochemical and biophysical research communications*, 334(4):1004–1013, 2005.

R.J. Read. Improved fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallographica Section A: Foundations of Crystallography*, 42 (3):140–149, 1986.

R.J. Read. Structure-factor probabilities for related structures. *Acta Crystallographica Section A: Foundations of Crystallography*, 46(11):900–912, 1990.

R.J. Read. Model phases: probabilities, bias and maps. *International Tables for Crystallography*, 2006.

R.L. Rich and D.G. Myszka. Survey of the year 2001 commercial optical biosensor literature. *Journal of Molecular Recognition*, 15(6):352–376, 2002.

J.S. Richardson. Schematic drawings of protein structures. *Methods in enzymology*, 115:359–380, 1985.

J.S. Richardson and D.C. Richardson. Amino acid preferences for specific locations at the ends of alpha helices. *Science*, 240(4859):1648–1652, 1988. ISSN 0036-8075.

Daniel Ricklin. *Surface Plasmon Resonance Applications in Drug Discovery*. PhD thesis, University Basel, 2005.

J.M. Rini, U. Schulze-Gahmen, and I.A. Wilson. Structural evidence for induced fit as a mechanism for antibody-antigen recognition. *Science*, 255(5047):959–965, 1992. ISSN 0036-8075.

R.J. Roberts and K. Murray. Restriction endonuclease. *Critical Reviews in Biochemistry and Molecular Biology*, 4(2):123–164, 1976. ISSN 1040-9238.

V.L. Roger, A.S. Go, D.M. Lloyd-Jones, R.J. Adams, J.D. Berry, T.M. Brown, M.R. Carnethon, S. Dai, G. de Simone, and E.S. Ford. Heart disease and stroke statistics - 2011 update1. *Circulation*, 123(4):e18–e209, 2011. ISSN 0009-7322.

G. Rosenbaum and K.C. Holmes. Synchrotron radiation as a source for x-ray diffraction. *Nature*, 230:434–437, 1971.

L.K. Roskos, C.G. Davis, and G.M. Schwab. The clinical pharmacology of therapeutic monoclonal antibodies. *Drug development research*, 61(3):108–120, 2004.

M.G. Rossmann. *The molecular replacement method*. Gordon and Breach, 1972.

M.G. Rossmann and D.M. Blow. The detection of sub-units within the crystallographic asymmetric unit. *Acta Crystallographica*, 15(1):24–31, 1962.

J. Rousseaux, R. Rousseaux-Prevost, and H. Bazin. Optimal conditions for the preparation of Fab and F(ab') 2 fragments from monoclonal IgG of different rat IgG subclasses. *Journal of immunological methods*, 64(1):141–146, 1983.

G. Rudenko, L. Henry, K. Henderson, K. Ichtchenko, M.S. Brown, J.L. Goldstein, and J. Deisenhofer. Structure of the ldl receptor extracellular domain at endosomal ph. *Science*, 298(5602):2353–2358, 2002. ISSN 0036-8075.

B. Rupp. *Biomolecular crystallography*. Garland Science, 2009.

J.E. Sadler. Biochemistry and genetics of von willebrand factor. *Annual review of biochemistry*, 67(1):395–424, 1998.

J.F. Sanchez, J. Lescar, V. Chazalet, A. Audfray, J. Gagnon, R. Alvarez, C. Breton, A. Imberty, and E.P. Mitchell. Biochemical and structural analysis of helix pomatia agglutinin. *Journal of Biological Chemistry*, 281(29):20171–20180, 2006.

C. Sander and R. Schneider. Database of homology-derived protein structures and the structural meaning of sequence alignment. *Proteins: Structure, Function, and Bioinformatics*, 9(1):56–68, 1991.

F. Sanger, A.R. Coulson, et al. A rapid method for determining sequences in dna by primed synthesis with dna polymerase. *Journal of molecular biology*, 94(3): 441–448, 1975.

F. Sanger, S. Nicklen, and A.R. Coulson. Dna sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences*, 74(12):5463–5467, 1977.

H. Schägger, G. Von Jagow, et al. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kda. *Analytical biochemistry*, 166(2):368, 1987.

O. Scharf, H. Golding, L.R. King, N. Eller, D. Frazier, B. Golding, and D.E. Scott. Immunoglobulin g3 from polyclonal human immunodeficiency virus (hiv) immune globulin is more potent than other subclasses in neutralizing hiv type 1. *Journal of virology*, 75(14):6558–6565, 2001.

J. Schmitt, H. Hess, and H.G. Stunnenberg. Affinity purification of histidine-tagged proteins. *Molecular biology reports*, 18(3):223–230, 1993.

R. Schneider, A. Sharma, and A. Rai. Introduction to molecular dynamics. *Computational Many-Particle Physics*, pages 3–40, 2008.

V. Schomaker and K.N. Trueblood. On the rigid-body motion of molecules in crystals. *Acta Crystallographica Section B: Structural Crystallography and Crystal Chemistry*, 24(1):63–76, 1968.

I.U. Schraufstatter, K. Trieu, L. Sikora, P. Sriramarao, and R. DiScipio. Complement c3a and c5a induce different signal transduction cascades in endothelial cells. *The Journal of Immunology*, 169(4):2102–2110, 2002.

H.W. Schroeder Jr and L. Cavacini. Structure and function of immunoglobulins. *Journal of Allergy and Clinical Immunology*, 125(2):S41–S52, 2010.

S. Schulman and N.R. Bijsterveld. Anticoagulants and their reversal. *Transfusion medicine reviews*, 21(1):37–48, 2007. ISSN 0887-7963.

N.G. Seidah, S. Benjannet, L. Wickham, J. Marcinkiewicz, S.B. Jasmin, S. Stifani, A. Basak, A. Prat, and M. ChrÃ©tien. The secretory proprotein convertase neural apoptosis-regulated convertase 1 (narc-1): liver regeneration and neuronal differentiation. *Proceedings of the National Academy of Sciences of the United States of America*, 100(3):928–, 2003. ISSN 0027-8424.

N.G. Seidah, G. Mayer, A. Zaid, E. Rousselet, N. Nassoury, S. Poirier, R. Essalmani, and A. Prat. The activation and physiological functions of the proprotein convertases. *The international journal of biochemistry & cell biology*, 40(6):1111–1125, 2008.

N.G. Seidah, S. Poirier, M. Denis, R. Parker, B. Miao, C. Mapelli, A. Prat, H. Wassef, J. Davignon, and K.A. Hajjar. Annexin a2 is a natural extrahepatic inhibitor of the pcsk9-induced ldl receptor degradation. *PloS one*, 7(7):e41865–, 2012. ISSN 1932-6203.

P.A. Sharp, B. Sugden, and J. Sambrook. Detection of two restriction endonuclease activities in haemophilus parainfluenzae using analytical agarose-ethidium bromide electrophoresis. *Biochemistry (Mosc).*, 12(16):3055–3063, 1973. ISSN 0006-2960.

B.W. Sigurskjold. Exact analysis of competition ligand binding by displacement isothermal titration calorimetry. *Analytical biochemistry*, 277(2):260–266, 2000.

M.K. Singh, S. Srivastava, GPS Raghava, and G.C. Varshney. Haptendb: a comprehensive database of haptens, carrier proteins and anti-hapten antibodies. *Bioinformatics*, 22(2):253–255, 2006.

E. Southern. Gel electrophoresis of restriction fragments. *Methods in enzymology*, 68: 152–, 1979. ISSN 0076-6879.

R.L. Stanfield, M. Takimoto-Kamimura, J.M. Rini, A.T. Profy, and I.A. Wilson. Major antigen-induced domain rearrangements in an antibody. *Structure*, 1(2):83–93, 1993. ISSN 0969-2126.

R.L. Stanfield, A. Zemla, I.A. Wilson, and B. Rupp. Antibody elbow angles are influenced by their light chain class. *Journal of Molecular Biology*, 357(5):1566–1574, 2006. ISSN 0022-2836.

J. Stangier. Clinical pharmacokinetics and pharmacodynamics of the oral direct thrombin inhibitor dabigatran etexilate. *Clinical pharmacokinetics*, 47(5):285–295, 2008. ISSN 0312-5963.

E.A. Stein, D. Gipe, J. Bergeron, D. Gaudet, R. Weiss, R. Dufour, R. Wu, and R. Pordy. Effect of a monoclonal antibody to pcsk9, regn727/sar236553, to reduce low-density lipoprotein cholesterol in patients with heterozygous familial hypercholesterolaemia on stable statin dose with or without ezetimibe therapy: a phase 2 randomised controlled trial. *The Lancet*, 2012. ISSN 0140-6736.

N. Stein. Chainsaw: a program for mutating pdb files used as templates in molecular replacement. *Journal of applied crystallography*, 41(3):641–643, 2008. ISSN 0021-8898.

E. Stenberg, B. Persson, H. Roos, and C. Urbaniczky. Quantitative determination of surface concentration of protein with surface plasmon resonance using radiolabeled proteins. *Journal of colloid and interface science*, 143(2):513–526, 1991.

A.W.J. Stuttle, M.J. Powling, J.M. Ritter, and R.M. Hardisty. Effects of a monoclonal antibody to glycoprotein iib/iiia (p256) and of enzymically derived fragments of p256 on human platelets. *Thrombosis and haemostasis*, 65(4):432–437, 1991.

D. Sullivan, A.G. Olsson, R. Scott, J.B. Kim, A. Xue, V. Gebski, S.M. Wasserman, and E.A. Stein. Effect of a monoclonal antibody to pcsk9 on low-density lipoprotein cholesterol levels in statin-intolerant patientsthe gauss randomized trialamg145 in statin-intolerant patients. *Journal of the American Medical Association*, 308(23): 2497–2506, 2012.

P.L. Surdo, M.J. Bottomley, A. Calzetta, E.C. Settembre, A. Cirillo, S. Pandit, Y.G. Ni, B. Hubbard, A. Sitlani, and A. Carfu. Mechanistic implications for ldl receptor degradation from the pcsk9/ldlr structure at neutral ph. *EMBO reports*, 12(12): 1300–1305, 2011.

BJ Sutton and HJ Gould. The human ige network. 1993.

I.J. Tickle, R.A. Laskowski, and D.S. Moss. Rfree and the rfree ratio. i. derivation of expected values of cross-validation residuals used in macromolecular least-squares refinement. *Acta Crystallographica Section D: Biological Crystallography*, 54(4): 547–557, 1998.

J.W. Tilley, L. Chen, D.C. Fry, S.D. Emerson, G.D. Powers, D. Biondi, T. Varnell, R. Trilles, R. Guthrie, F. Mennona, et al. Identification of a small molecule inhibitor of the il-2/il-2ralpha receptor interaction which binds to il-2. *Journal of the American Chemical Society*, 119(32):7589–7590, 1997.

A. Tiselius and E.A. Kabat. An electrophoretic study of immune sera and purified antibody preparations. *The Journal of experimental medicine*, 69(1):119–131, 1939.

J.B. Treweek and K.D. Janda. An antidote for acute cocaine toxicity. *Molecular pharmaceutics*, 9(4):969–978, 2012.

T. Turbadar. Complete absorption of light by thin metal films. *Proceedings of the Physical Society*, 73:40–, 1959.

T. Unge. Crystallization methods. *International University Line, La Jolla, Calif*, 1999.

A.A. Vagin, G.N. Murshudov, and B.V. Strokopytov. Blanc: the program suite for protein crystallography. *Journal of applied crystallography*, 31(1):98–102, 1998.

J. Van Ryn, J. Stangier, S. Haertter, K.H. Liesenfeld, W. Wienen, M. Feuring, and A. Clemens. Dabigatran etexilate - a novel, reversible, oral direct thrombin inhibitor: Interpretation of coagulation assays and reversal of anticoagulant activity. *Thromb Haemost*, 103(6):1116–1127, 2010.

B.K. Van Weeman and A. Schuurs. Immunoassay using antigen-enzyme conjugates. *FEBS Letter*, 15:232–235, 1971.

M. Varret, M. Abifadel, J.P. Rabes, and C. Boileau. Genetic heterogeneity of autosomal dominant hypercholesterolemia. *Clinical genetics*, 73(1):1–13, 2008. ISSN 1399-0004.

J. Venkatraman, S.C. Shankaramma, and P. Balaram. Design of folded peptides. *Chemical reviews*, 101(10):3131–3152, 2001. ISSN 0009-2665.

B. Vogelstein and D. Gillespie. Preparative and analytical purification of dna from agarose. *Proceedings of the National Academy of Sciences*, 76(2):615–, 1979. ISSN 0027-8424.

C. Vonrhein, C. Flensburg, P. Keller, A. Sharff, O. Smart, W. Paciorek, T. Womack, and G. Bricogne. Data processing and analysis with the autoproc toolbox. *Acta Crystallographica Section D: Biological Crystallography*, 67(Pt 4):293–302, April 2011. ISSN 1399-0047.

A.L. Waldo, R.C. Becker, V.F. Tapson, and K.J. Colgan. Hospitalized patients with atrial fibrillation and a high risk of stroke are not being provided with adequate anticoagulation. *Journal of the American College of Cardiology*, 46(9):1729–, 2005.

W. Wienen, J. Stassen, H. Priepke, U.J. Ries, and N. Hauel. In-vitro profile and ex-vivo anticoagulant activity of the direct thrombin inhibitor dabigatran and its orally active prodrug, dabigatran etexilate. *Thrombosis and Haemostasis*, 98(1):155–, 2007. ISSN 0340-6245.

M. Wilchek, E.A. Bayer, et al. The avidin-biotin complex in bioanalytical applications. *Analytical biochemistry*, 171(1):1, 1988.

A.F. Williams and A.N. Barclay. The immunoglobulin superfamily-domains for cell surface recognition. *Annual review of immunology*, 6(1):381–405, 1988.

A.J.C. Wilson. The probability distribution of x-ray intensities. *Acta Crystallographica*, 2(5):318–321, 1949.

I.A. Wilson and R.L. Stanfield. Antibody-antigen interactions: new structures and new conformational changes. *Current opinion in structural biology*, 4(6):857–867, 1994. ISSN 0959-440X.

M.D. Winn, M.N. Isupov, and G.N. Murshudov. Use of tls parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallographica Section D: Biological Crystallography*, 57(1):122–133, 2001.

J.M. Woof and J. Mestecky. Mucosal immunoglobulins. *Immunological reviews*, 206 (1):64–82, 2005.

T. Yamamoto, C. Lu, and R.O. Ryan. A two-step binding model of pcsk9 interaction with the low density lipoprotein receptor. *Journal of Biological Chemistry*, 286(7): 5464–5470, February 2011. ISSN 1083-351X.

D.W. Zhang, T.A. Lagace, R. Garuti, Z. Zhao, M. McDonald, J.D. Horton, J.C. Cohen, and H.H. Hobbs. Binding of proprotein convertase subtilisin/kexin type 9 to epidermal growth factor-like repeat a of low density lipoprotein receptor decreases receptor recycling and increases degradation. *Journal of Biological Chemistry*, 282 (25):18602–, 2007. ISSN 0021-9258.

D.W. Zhang, R. Garuti, W.J. Tang, J.C. Cohen, and H.H. Hobbs. Structural requirements for pcsk9-mediated degradation of the low-density lipoprotein receptor. *Proceedings of the National Academy of Sciences*, 105(35):13045–, 2008.

Z. Zhao and P. Michaely. The epidermal growth factor homology domain of the ldl receptor drives lipoprotein release through an allosteric mechanism involving h190, h562, and h586. *Journal of Biological Chemistry*, 283(39):26528–26537, 2008. ISSN 0021-9258.

A. Zhou, S. Martin, G. Lipkind, J. LaMendola, and D.F. Steiner. Regulatory roles of the p domain of the subtilisin-like prohormone convertases. *Journal of Biological Chemistry*, 273(18):11107–11114, 1998. ISSN 0021-9258.

A Appendix

Crystallographic tables

	aDabi-Fab1:dabigatran	aDabi-Fab1
Data collection		
Wavelength [Å]	0.9100	1.0000
Space group	$C222_1$	$C2$
Cell parameters		
a [Å]	74.16	113.13
b [Å]	223.82	78.22
c [Å]	157.21	63.43
α [deg]	90.00	90.00
β [deg]	90.00	89.91
γ [deg]	90.00	90.00
Resolution range [Å]	157.21 – 1.71 (1.80 – 1.71)	64.34 – 1.89 (1.89 – 1.94)
No. observations	946382 (142092)	134967 (18637)
No. unique observations	141489 (20492)	75791 (10891)
R_{meas} [%]	5.5 (55.2)	8.9 (60.5)
Average I/σI	21.6 (3.7)	10.7 (2.3)
Completeness [%]	99.9 (100)	99.6 (100.0)
Multiplicity	6.7 (6.9)	3.3 (3.3)
Refinement		
Matthews coefficient	3.4	2.9
Solvent content	63.79	58.03
Molecules in ASU	2	1
No. reflections	140590	44143
R_{work}	16.7	17.6
R_{free}	18.2	19.7
No. protein atoms	6796	3409
No. dabigatran atoms	70	-
No. water molecules	1306	366
Average B values [Å^2]		
Antibody	29.5	30.5
Water molecules	45.2	44.5
Dabigatran	25.2	-
Ramachandran statistics [%]		
Most favoured	98.5	98.4
Allowed	1.4	1.4
Disallowed	0.1	0.2
rmsd from ideality		
Bond lengths [Å]	0.008	0.008
Bond angles [deg]	1.04	1.05

	aDabi-Fab2	aDabi-Fab2: Dabigatran 1	aDabi-Fab2: Dabigatran 2	aDabi-Fab3	aDabi-Fab3: Dabigatran
Data collection					
Wavelength [Å]	1.00002	1.00002	1.00002	1.00002	1.00002
Space group	$P1$	$P2_1$	$P2_12_12_1$	$P6_1$	$P2_12_12_1$
Cell parameters					
a [Å]	55.93	51.81	48.20	103.69	59.97
b [Å]	56.15	128.92	59.74	103.69	78.39
c [Å]	81.50	60.26	127.69	84.92	87.67
α [deg]	87.38	90.00	90.00	90.00	90.00
β [deg]	82.97	92.30	90.00	90.00	90.00
γ [deg]	65.59	90.00	90.00	120.00	90.00
Resolution range [Å]	51.13 – 1.76	128.92 – 1.90	127.69 – 2.21	51.85 – 2.09	87.67 – 2.16
No. observations	154351 (22607)	202218 (29154)	119943 (17227)	309519 (46252)	144508 (17047)
No. unique observations	85149 (12315)	61726 (9015)	19263 (2767)	30655 (4455)	22495 (3019)
R_{meas} [%]	10.7 (44.5)	12.1 (60.1)	15.8 (63.8)	10.6 (62.7)	14.0 (59.4)
Average I/σI	7.5 (2.3)	9.6 (2.4)	9.8 (3.2)	15.8 (3.4)	11.0 (3.2)
Completeness [%]	96.2 (95.0)	99.4 (99.8)	99.9 (99.8)	100 (100)	98.7 (93.5)
Multiplicity	1.8 (1.8)	3.3 (3.3)	6.2 (6.2)	10.1 (10.4)	6.4 (5.6)
Refinement					
Matthews coefficient	2.40	2.08	1.90	2.76	2.14
Solvent content	48.70	40.98	35.45	55.54	42.57
Molecules in ASU	2	2	1	1	1
No. reflections	85139	61647	19095	30575	22428
R_{work}	17.81	18.11	17.43	18.13	16.97
R_{free}	19.82	21.46	23.41	22.11	21.98
No. protein atoms	6071	6733	3356	3330	3354
No. dabigatran atoms	-	70	35	-	35
No. water molecules	827	855	250	326	321
Average B values [Å²]					
Antibody	20.54	22.74	28.08	39.81	26.09
Water molecules	31.51	30.68	34.31	47.32	35.67
Dabigatran	-	27.80	35.46	-	36.59
Ramachandran statistics [%]					
Most favoured	97.8	97.5	96.3	96.3	97.9
Allowed	2.1	2.4	3.4	3.7	2.1
Disallowed	0.1	0.1	0.2	0.0	0.0
rmsd from ideality					
Bond lengths [Å]	0.008	0.008	0.008	0.008	0.008
Bond angles [deg]	1.03	1.07	1.10	1.07	1.08

	aDabi-Fab1:C1	aDabi-Fab1:C3	aDabi-Fab1:C4	aDabi-Fab1:C5	aDabi-Fab1:C6
Data collection					
Wavelength [Å]	1.00004	1.00004	1.00004	1.00004	1.00004
Space group	$C222_1$	$C222_1$	$C222_1$	$C222_1$	$C222_1$
Cell parameters					
a [Å]	74.34	74.25	74.50	74.16	74.30
b [Å]	224.24	224.07	224.47	224.44	224.09
c [Å]	157.60	157.45	157.34	157.70	157.33
α [deg]	90.00	90.00	90.00	90.00	90.00
β [deg]	90.00	90.00	90.00	90.00	90.00
γ [deg]	90.00	90.00	90.00	90.00	90.00
Resolution range [Å]	157.60 - 1.98	386.11 - 3.28	157.3 - 2.08	157.7 - 2.69	157.3 - 1.74
No. observations	617154	666374	533102	269571	888209
No. unique observations	92311	97850	79375	36276	133581
R:meas [%]	6.7 (66.4)	7.5 (55.2)	8.7 (49.1)	8.7 (53.0)	4.5 (46.4)
Average I/σI	18.8 (3.3)	18.2 (3.9)	15.34 (3.39)	20.17 (3.83)	21.52 (3.41)
Completeness [%]	100.0 (100.0)	97.6 (96.8)	99.98 (99.99)	98.04 (99.61)	99.90 (99.79)
Multiplicity	6.7 (7.0)	6.8 (7.0)	6.7 (6.6)	7.3 (7.3)	6.6 (6.6)
Refinement					
Matthews coefficient	3.58	3.68	3.64	3.61	3.66
Solvent content	65.67	66.59	66.24	65.91	66.39
Molecules in ASU	2	2	2	2	2
No. reflections	85139	97850	79375	36276	133581
R_{work}	17.9	17.7	17.6	21.8	17.7
R_{free}	20.0	19.7	20.5	26.0	19.4
No. protein atoms	6836	6652	6721	6787	6692
No. compound atoms	82	74	74	72	86
No. water molecules	868	912	951	301	974
Average B values [Å²]					
Antibody	42.70	33.50	33.00	50.30	32.50
Water molecules	51.50	46.10	44.70	42.80	44.70
Ramachandran statistics [%]					
Most favoured	98.3	98.4	97.1	97.1	98.7
Allowed	1.7	1.6	2.3	2.5	1.3
Disallowed	0.0	0.0	0.6	0.3	0.0
rmsd from ideality					
Bond lengths [Å]	0.008	0.008	0.008	0.008	0.008
Bond angles [deg]	1.05	1.03	1.08	1.10	1.07

	aDabi-Fab1:C7	aDabi-Fab1:C8	aDabi-Fab1:C9	aDabi-Fab1:C10	aDabi-Fab1:C11
Data collection					
Wavelength [Å]	1.00004	1.00004	1.00004	1.00004	1.00004
Space group	$C222_1$	$P2_1$	$C222_1$	$C222_1$	$C222_1$
Cell parameters					
a [Å]	74.11	54.86	74.13	74.26	74.08
b [Å]	223.23	91.97	223.306	223.94	223.73
c [Å]	156.86	89.67	156.852	157.22	157.08
α [deg]	90.00	90.00	90.00	90.00	90.00
β [deg]	90.00	96.89	90.00	90.00	90.00
γ [deg]	90.00	90.00	90.00	90.00	90.00
Resolution range [Å]	156.9 - 1.88	91.97 - 1.52	156.8 - 2.42	157.2 - 1.8	157.1 - 1.95
No. observations	705843	454861	369788	804866	679861
No. unique observations	105436	134224	50070	121102	94951
R_{meas} [%]	5.0 (49.3)	4.0 (41.7)	9.8 (55.3)	5.2 (48.8)	5.4 (48.1)
Average I/σI	21.57 (3.49)	13.51 (2.55)	17.70 (3.62)	20.29 (3.35)	23.46 (3.57)
Completeness [%]	99.97 (99.99)	99.00 (98.77)	100.00 (100.00)	99.89 (99.98)	99.97 (99.97)
Multiplicity	6.7 (6.9)	3.4 (3.5)	7.3 (7.2)	6.6 (6.5)	7.2 (6.7)
Refinement					
Matthews coefficient	3.63	2.49	3.61	3.58	3.59
Solvent content	66.12	50.59	65.91	65.67	31.56
Molecules in ASU	2	2	2	2	2
No. reflections	105436	134224	50070	121102	94951
R_{work}	17.5	18.8	16.0	17.7	17.3
R_{free}	19.5	20.7	20.7	19.6	19.9
No. protein atoms	6692	6715	6788	6836	6813
No. compound atoms	84	76	70	78	72
No. water molecules	944	757	1017	1022	1210
Average B values [Å²]					
Antibody	34.90	27.50	34.50	31.90	35.70
Water molecules	46.60	36.60	45.40	44.90	48.50
Ramachandran statistics [%]					
Most favoured	98.4	97.5	97.7	98.6	98.3
Allowed	1.6	2.5	2.3	1.4	1.6
Disallowed	0.0	0.0	0.0	0.0	0.1
rmsd from ideality					
Bond lengths [Å]	0.008	0.008	0.008	0.008	0.008
Bond angles [deg]	1.06	1.05	1.03	1.05	1.05

	aPCSK9-Fab1	aPCSK9-Fab1: peptide	aPCSK9-Fab2	aPCSK9-Fab2 ternary complex	aPCSK9-Fab3
Data collection					
Wavelength [Å]	1.00004	0.90000	1.00004	0.90000	0.90000
Space group	*C*2	*P*1	$P2_12_12$	*C*2	$P2_12_12_1$
Cell parameters					
a [Å]	93.22	40.00	92.06	262.19	91.83
b [Å]	87.98	42.74	62.34	138.65	100.01
c [Å]	69.83	56.13	84.01	69.54	103.70
α [deg]	90.00	92.82	90.00	90.00	90.00
β [deg]	123.47	97.67	90.00	102.87	90.00
γ [deg]	90.00	90.35	90.00	90	90.00
Resolution range [Å]	58.26 – 2.16	42.68 – 1.42	92.06 – 2.19	127.80 – 3.22	103.7 - 1.95
No. observations	91766 (12816)	172989 (23587)	180730 (25850)	133892 (19686)	467343(50248)
No. unique observations	25206 (3647)	93961 (12996)	25422 (3666)	38982 (5677)	70073(10187)
R_{meas} [%]	11.1 (70.6)	5.9 (30.2)	15.8 (57.8)	22.2 (70.9)	4.3 (51.5)
Average I/σI	11.8 (2.4)	10.1 (3.3)	13.4 (4.1)	6.6 (2.1)	24.3 (3.0)
Completeness [%]	99.9 (99.9)	89.8 (85.0)	100 (100)	99.3 (99.7)	99.9 (99.8)
Multiplicity	3.6 (3.5)	1.8 (1.0)	7.1 (7.1)	3.4 (3.5)	6.6 (6.4)
Refinement					
Matthews coefficient	2.49	1.98	2.51	3.67	2.78
Solvent content	50.60	37.88	51.05	66.48	55.77
Molecules in ASU	1	1	1	1	2
No. reflections	20925	67598	25209	38968	70073
R_{work}	19.91	19.35	17.16	16.74	18.96
R_{free}	24.92	21.30	24.51	22.64	21.57
No. protein atoms	3209	3356	3336	10963	6389
No. water molecules	304	336	513	282	588
Average B values [Å^2]					
Antibody	35.03	19.61	28.08	81.56	47.30
Water molecules	38.46	27.88	34.31	39.19	50.60
Ramachandran statistics [%]					
Most favoured	95.2	96.5	97.5	94.5	97.8
Allowed	4.3	3.3	2.5	4.6	2.0
Disallowed	0.5	0.2	0.0	0.9	0.2
rmsd from ideality					
Bond lengths [Å]	0.008	0.008	0.008	0.008	0.008
Bond angles [deg]	1.17	1.09	1.10	1.09	1.03

Acknowledgments

First of all I would like to thank Univ.-Prof. Dr. Michael Groll from the Chair of Biochemistry at the TU München for accepting the supervision of my external Ph.D. thesis. Due to his openhearted and supportive character, I always enjoyed visiting the chair for our status updates and any correspondence regarding all the bureaucratic issues.

Thanks to Univ.-Prof. Dr. Thomas Kiefhaber from the Chair of Biophysical Chemistry at the TU München for being second examiner of my Ph.D. thesis.

I want to thank Nerea Gallastegui and Philipp Baer for supervising the practical courses for biochemistry students with me. It was always a lot of fun and definitely worth traveling to Munich for 3 weeks, once a year.

Thanks to Dr. John Park, Dr. Peter Maier and Dr. Tobias Litzenburger for introducing me to the NBE unit and for giving very valuable scientific input at our regular meetings. Thanks to John for the evening session at the IBM, to Tobias for helping me with the submission to blood and to Peter for struggling with the PCSK9 project with me.

I'm very grateful to Michael Ritter and Nikolai Roosz for welcoming me in their lab with open arms. It was fun to refresh my cell culture techniques under your guidance and to discuss the best ways of getting things done. Thanks for explaining the KinExA technology and for always supporting me with any protein in times of need.

Many thanks to Martina Hafner and Karolin Schwarz. Working with you on the PCSK9 project was always fun and the results were quite promising. I always felt welcome in your lab and learned a lot.

Thanks to Dr. Daniel Seeliger for continuously supporting me in questions regarding molecular dynamics simulations and for helping me with the MD evaluation.

I'd like to thank Kit Briant for the Oxford United scarf and for welcoming me and Jo at Manchester. Thanks to Donia Al Saffar for being a very pleasant office neighbor and for sharing her black humor with me. Thanks to Danielle Owen for discussions in the lab, especially for lessons on the Scottish play.

Thanks to Dr. Margit Bauer for her guidance in questions of refolding and the members of PEP1 for their cooperation. Very special thanks to Ingo Hinz who introduced me to the Biacore SPR instrument. Special thanks for sharing your thoughts on high-quality TV programs and the Bundesliga.

I want to thank the colleagues from the NMR Lab. I always enjoyed the discussions with you. Markus, keep your head up, better days for the KSC will come.

Thanks to Dr. Gisela Schnapp and her lab for continuously expressing PCSK9 for

my studies. Dr. Schnapp always supported me in issues concerning molecular biology and protein purification (and hiking). Thanks to Yvette Hoevels for introducing me to ITC, to Jessica Bretzel for maintaining “my own” FPLC, to Heike Rapp for all kinds of support (not only for exchanging water in the buckets) and to Adelheid Loehle for the extensive production of PCSK9 and discussions on football in general (and Lothar Matthäus in detail). Instead of acknowledging every single aspect, I just want to say DANKE to all of you.

Many thanks to Nadja Stadler, Anita Bloching, Angela Schmidt and Heidi Roth for the crystallization service. Your help, especially with the awful Phoenix was invaluable. Your hints on crystallization helped to push my projects forwards. Thanks for crystallizing late on Friday evenings and for fulfilling extraordinary suggestions regarding crystallization screens.

Thanks to all the crystallographers for always helping me with any kind of crystallographic issue for the last three years. Thanks to Dr. Stefan Hörer for supplying me with constant amounts of Bergkäse and for revising a part of the manuscript for this thesis. I want to thank Dr. Dennis Fiegen for constructive discussions on X-ray crystallography and for accompanying my first SLS trip. It was a pleasure talking about things not correlated to work. Thanks to Dr. Alex Pautsch for sharing a significant part of his science fiction and fantasy library with me. Especially the light-weight tent was invaluable on the West Highland Way. I want to acknowledge Dr. Dirk Reinert who was willing to share his office with me during the first 2 years and also helping me with any questions concerning X-ray software, shell scripts and chewing gum.

Finally I’d like to thank Dr. Herbert Nar for supervising my Ph.D. thesis as head of the Structural Research Group at Boehringer Ingelheim in Biberach. His scientific guidance and discussions surely led me the right way during more than 3 years of very pleasant work (it actually never felt like work, rather like a really interesting adventure). Special thanks for allowing me to work on my own, while always supporting me with the right and motivating suggestions when independent work led to a dead end. I can hardly say how much I enjoyed these years. Sometimes gratitude is more than some written lines . . .

Declaration

I, Felix Schiele, declare that this thesis is my own work and has not been submitted in any form at any other university or institute. Information derived from the published and unpublished work of others has been acknowledged in the text and a list of references is given in the bibliography. Parts of this thesis will be published in scientific journals.

Felix Schiele

Berlin, 1st May, 2013 .....