Substrate specificity of Glycine Oxidase
and
protein interaction specificity of the neuronal cell adhesion molecule TAG-1

Dissertation
zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften (Dr. rer. nat.)

Universität Konstanz
Mathematisch-Naturwissenschaftliche Sektion
Fachbereich Biologie
Lehrstuhl für Biophysik
Prof. Dr. Wolfram Welte

vorgelegt von Diplom Biologe Mario Mörtl

Referenten:
Prof. Dr. Wolfram Welte (Konstanz)
Prof. Dr. Peter Sonderegger (Zürich)

Tag der mündlichen Prüfung, 8.12.2006

Index

Zusammenfassung

Die Entwicklung des Nervensystems erfordert unterschiedliche Wechselwirkungen des wachsenden Axones mit seiner Umgebung, damit die Neuronen ihre oft weit entfernten Ziele erreichen können, um dort mit anderen Zellen Kontaktstellen auszubilden. Dabei finden die Axone den richtigen Weg, indem sie sich an molekularen Markierungen orientieren, die der Wachstumskegel der Axone erkennen kann. Solche Moleküle werden oft von den Axonen auch selbst sekretiert. Die größte Gruppe stellen dabei die in der Membran verankerten neuronalen Zelladhäsionsmoleküle dar, besonders die aus der Immunglobulinfamilie. Die Zelladhäsionsmoleküle dieser Gruppe sind oft Ligand und Rezeptor in einem und nehmen die jeweilige Rolle abhängig vom Entwicklungsgrad des Nervensystems und ihrer aktuellen Umgebung ein. Um die zugrunde liegenden molekularen Interaktionsmechanismen zu verstehen, wurden in dieser Arbeit folgende sechs Vertreter dieser Familie untersucht: menschliches TAG-1, menschliches L1 und seine Huhn- und Goldfisch-Homologen NgCAM (Neuronglia cell adhesion molecule) und E587-antigen, NrCAM vom Huhn (NgCAM related cell adhesion molecule) und F11 (RAR2/CNTN5) vom Huhn. Für alle untersuchten Moleküle wurden Expressionsvektoren der vier aminoterminalen Immunoglobulindomänen konstruiert. Frühere Untersuchungen ließen darauf schließen, dass alle untersuchten Moleküle eine zueinander ähnliche Tertiär-Struktur aufweisen: Ein kompaktes Konglomerat, in der die vier Immunoglobulindomänen U-förmig angeordnet sind. Alle Konstrukte konnten in *E. coli* als unlösliche Inclusion Bodies exprimiert werden. Der experimentell schwierigste Schritt bestand in der oxidativen *in vitro* Rückfaltung. Zwei der untersuchten Proteine (TAG-1 und E587-antigen) konnten erfolgreich rückgefaltet und löslich aufgereinigt werden. Für TAG-1 gelang die Herstellung von Kristallen, und mit Hilfe der Röntgenstrukturanalyse konnte ein dreidimensionales TAG-1 Modell gewonnen werden. Anhand der spezifischen Anordnung der TAG-1-Moleküle im Kristallgitter, wurde ein alternatives Interaktionsmodell zur Erklärung der homophilen *trans* Interaktion von TAG-1 entwickelt. Das hier vorgestellte Interaktionsmodell beruht auf der *Trans*-Interaktion der vier aminoterminalen Immunglobulindomänen zweier TAG-1 Moleküle, indem ein stabiles Dimer entsteht. Damit unterscheidet sich das neue Modell vom molekularen Reißverschlusses, welches von der Axonin-1-Struktur abgeleitet wurde (FREIGANG *ET AL.*, 2000), denn dort ensteht mit dem Reißverschluss eine Art Superkomplex, dessen Größe von der Anzahl der beteiligten Moleküle (Reißverschlusszähne) abhängt. Interessanterweise stützen jene Experimente, die schon das Reißverschlußmodell im Falle von Axonin-1 stützten, auch das Dimer-Interaktions-Modell, das hier vorgeschlagen wird, da die für die *Trans*-Interaktion verantwortlichen Reste für beide Modelle übereinstimmen. Das aus der TAG-1 Struktur abgeleitete Modell kann

zudem einfach zu einem Modell erweitert werden, das die von KUNZ *ET AL.*, 2002 gezeigte, homophile *Cis*-Interaktion durch die Fibronectin-III-Domänen mit berücksichtigt.

Unabhängig von der Gruppe Ealick (SETTEMBRE *ET AL.*, 2003) wurde die 1,8 Å Komplexstruktur der Glyzinoxidase von *B. subtilis* mit dem Inhibitor Glykolat gelöst. Die Struktur wurde mit der Methode des multiplen isomorphen Ersatzes mit Hilfe zweier Schweratomderivate gelöst. Die Glycinoxydase ist ein Homotertamer, was bisher einzigartig ist, für ein Mitglied der GR_2 Familie der Glutathionreduktasen. Die Glyzinoxidase soll eine entscheidende Rolle in der Thiaminbiosynthese spielen (SETTEMBRE *ET AL.*, 2003). Diese Rolle wird von der Kristallstruktur nicht unterstützt, da das Reaktionzentrum durch einen Subtrattunnel direkt mit dem Lösungsmittel in Verbindung steht. Der zweite Schritt der Thiaminbiosynthese, die Reaktion von Thiocarboxylat mit dem Iminprodukt muss jedoch unter Wasserausschluss erfolgen, da das Iminprodukt sonst hydrolisiert würde. Ein von der Glyzinoxidase bereitgestellter Mechanismus, der den Kontakt des Iminproduktes zum Wasser verhindert, kann aus der Röntgenstruktur nicht abgeleitet werden. Es konnte in dieser Arbeit auch gezeigt werden, dass die enzymatische Aktivität nicht durch die Anordnung als Homotetramer beeinflusst wird, da kein allosterischer Effekt durch Phosphat, Thiamin oder Thiaminpyrophosphat auf den Glyzinumsatz beobachtet wurde.

Summary

Cell adhesion molecules interact among each other during the development of the nervous system and guide axons to reach their targets. Axonal growth during the embryonic development requires molecular markers, so called guidance cues, which can be detected and secreted by the growth cone of commissural axons. Factors, which led to correct path finding are membrane bound neuronal cell adhesion molecules, as well as soluble components, and extracellular matrix components. Members of the large immunoglobulin family of neuronal cell adhesion molecules are involved in those developmental processes. Molecules of this family act as ligands and receptors during path finding, depending on the developmental status of the growing axon and on its location on the way to the target. To unravel the underlying molecular principle of interaction, six different molecules were investigated here: Human TAG-1, which is the ortholog of chicken Axonin-1, human L1 and its chicken and goldfish homologues NgCAM (Neuronglia cell adhesion molecule) and E587-antigen, chicken NrCAM (NgCAM related cell adhesion molecule), and chicken F11 (RAR2/CNTN5). For all targets, expression constructs of the four N-terminal immunoglobulin domains were constructed. All target proteins are supposed to adopt a similar tertiary structure, namely a compact conglomerate, where the four Ig domains are arranged in an U-shaped manner, with a strong (180 degrees) bent between the second and the third domain. Expression in *E. coli* was possible for all target molecules as insoluble inclusion bodies. The next experimental step, the oxidative *in vitro* refolding was very difficult, but finally TAG-1 and E587-antigen could be refolded and obtained in soluble form, suitable for crystallization experiments. In the case of TAG-1, the crystal structure was solved, and an alternative mechanism for homophilic *trans* interaction of TAG-1 was proposed, derived from the lattice contacts, found in the TAG-1 crystal. The interaction model is characterized by homodimerization of two TAG-1 molecules. The interaction interface is formed mainly by residues located on second immunoglobulin domain. The new interaction model is different from the zipper model proposed earlier on the basis of the Axonin-1 crystal structure (Freigang *et al.*, 2000), as the latter resulted in a large complex, where the size is determined by the number of interacting molecules (equivalent to the teeth of the zipper). Interestingly those experiments which supported the zipper model in the case of Axonin-1, do also support the dimer interaction model proposed here. In addition, the TAG-1 derived model can be easily expanded to a more complex model, where the homophilic *cis* interaction of the fourth FnIII domain (Kunz *et al.*, 2002) is considered too.

Independently from the group of Ealick (Settembre *et al.*, 2003), glycine oxidase of *B. subtilis* was structurally characterized with a resolution of 1.8 Å in complex with the inhibitor glycolate. The structure was solved, using the multiple isomorphous replacement method

using two heavy atom derivatives. Glycine oxidase is a tetrameric enzyme, which is unique for the GR_2 subfamily of glutathion reductases, and was supposed to play a crucial role in thiamine biosynthesis. This role of GO is not supported by the crystal structure, because the active site of GO is in direct contact with the bulk solvent via a small substrate channel. But the second step in the thiamine biosynthesis pathway, the reaction of thiocarboxylate with the imine product of glycine oxidase, has to be catalyzed in the absence of water, otherwise the imine product would be hydrolyzed quickly. A mechanism provided by the glycine oxidase, suitable to prevent the imine product from water contact, until the following enzyme of the thiazole formation pathway (ThiS) has acted, could not be derived from the structural data. In addition, this work could show, that the enzymatic properties are not affected by the homotetrameric quaternary structure, as no allosteric effect on glycine oxidase activity could be observed by phosphate, thiamine, or thiamine pyrophosophate.

Abbreviations

ASU	Asymmetric unit
AxCAM	Axonal cell adhesion molecule
B	Buried surface
CV	column volume(s)
DAAO	D-amino acid oxidase
F_{ab}	Antigen binding fragment of an antibody, composed of the light chain and the variable and the first constant domain of the heavy chain
f_{bu}	Fraction of fully buried atoms
FnIII-domain	Fibronectin-III-like domain
$f_{np}B$	Non-polar interface area
GO	Glycine oxidase
gor	Glutathion-Oxoreductase
GR	Glutathione reductase
IEF	Isoelectric focusing
Ig-domain	Immunoglobulin like domain
IgSF	Immunoglobulin superfamily
IMAC	Immobilized metal affinity chromatography
IPTG	Isopropyl-β-D-thiogalactopyranosid
lacI	Lac-repressor
MSOX	Monomeric sarcosine oxidase
PCR	Polymerase chain reaction
pkDAAO	Pig kidney D-amino acid oxidase
rgDAAO	*Rhodotorula gracilis* D-amino acid oxidase
RMS	Root mean square
RP	Residue propensity score
S_C	Shape complementarity value
trxB	Thioredoxin reductase

A. Structure-function correlation of Glycine oxidase from *Bacillus subtilis*

This work was published in:

***MÖRTL, M**., DIEDERICHS, K., WELTE, W., MOLLA, G., MOTTERAN, L., ANDRIOLO, G., PILONE, M. S. & POLLEGIONI, L. (2004) Structure-function correlation in glycine oxidase from Bacillus subtilis. J Biol Chem, **279**, 29718-27.*

A.1. Abstract

Structure-function relationships of the flavoprotein glycine oxidase (GO), which recently has been proposed as the first enzyme in the biosynthesis of thiamine in *Bacillus subtilis,* has been investigated by a combination of structural and functional studies. The structure of the GO-glycolate complex was determined at 1.8 Å, a resolution at which a sketch of the residues involved in FAD binding and in substrate interaction can be depicted: GO can be considered a member of the amine oxidase class of flavoproteins, such as D-amino acid oxidase and monomeric sarcosine oxidase. With the obtained model of GO the monomer-monomer interactions can be analyzed in detail, thus explaining the structural basis of the stable tetrameric oligomerization state of GO, which is unique for the GR_2 subfamily of flavoxidases. On the other hand, the 3D structure of GO and the functional experiments do not provide the functional significance for such an oligomerization state; GO does not show an allosteric behavior. The results do not clarify the metabolic role of this enzyme in *B. subtilis;* the broad substrate specificity cannot be correlated with the inferred function in thiamine biosynthesis and the structure does not show how GO could interact with ThiS, the following enzyme in thiamine biosynthesis. However, they do let a general catabolic role of this enzyme on primary or secondary amines to be excluded because the expression of GO is not inducible by glycine, sarcosine or D-alanine as carbon or nitrogen sources.

A.2. Introduction

Glycine oxidase (GO, EC 1.4.3.19) is a flavoprotein consisting of four identical subunits (369 residues each) and containing one molecule of non-covalently bound FAD per 42 kDa protein molecule (JOB *ET AL.*, 2002A; JOB *ET AL.*, 2002B). GO catalyzes a reaction similar to that of D-amino acid oxidase (DAAO, EC 1.4.3.3), a paradigm of the dehydrogenase-oxidase class of flavoproteins (for a recent review see PILONE, 2000). Both enzymes catalyze the oxidative deamination of amino acids to yield the corresponding α-imino acids and, after hydrolysis, α-keto acids, ammonia (or primary amines), and hydrogen peroxide. Both enzymes show a high pK_a for flavin N-3H ionization, do not bind covalently the FAD cofactor and react readily with sulfite (JOB *ET AL.*, 2002A; JOB *ET AL.*, 2002B; PILONE, 2000), but they differ in substrate specificity. In addition to neutral D-amino acids (e.g. D-alanine, D-proline, etc., which are also good substrates of DAAO), GO catalyzes the oxidation of primary and secondary amines (e.g. glycine, sarcosine, etc.) partially sharing the substrate specificity with monomeric sarcosine oxidase (MSOX, EC 1.5.3.1), an enzyme that catalyzes the oxidative demethylation of sarcosine to yield glycine, formaldehyde and hydrogen peroxide (WAGNER AND SCHUMAN

JORNS, 1997). According to investigations of the substrate specificity and of the binding properties, the GO active site seems to preferentially accommodate amines of a small size, such as glycine and sarcosine (JOB *ET AL.*, 2002A; JOB *ET AL.*, 2002B). GO follows a ternary complex sequential mechanism with glycine, sarcosine and D-proline as substrates in which the rate of product dissociation from the re-oxidized enzyme form represents the rate-limiting step (MOLLA *ET AL.*, 2003). Such a kinetic mechanism is similar to that determined for mammalian DAAO on neutral D-amino acids and for the MSOX on L-proline (PORTER *ET AL.*, 1977; WAGNER AND JORNS, 2000); the main difference is represented by the observed reversibility of the GO reductive half-reaction.

Taken together, however, these results do not really clarify the function of GO in *B. subtilis*; they only outline a general catabolic role for GO (JOB *ET AL.*, 2002A; JOB *ET AL.*, 2002B). Recent work proposed GO as the first enzyme in the biosynthesis of the thiazole moiety of thiamine pyrophosphate cofactor in *B. subtilis* (SETTEMBRE *ET AL.*, 2003). According to this hypothesis, GO catalyzes the oxidation of glycine to give the imine product that would be trapped with the thiocarboxylate intermediate bound to the following enzyme of the pathway (ThiS). In such a mechanism, the nucleophilic addition might occur at the active site of GO to avoid hydrolysis of the imino product. The known 2.3 Å resolution structure of GO (SETTEMBRE *ET AL.*, 2003) does not provide direct evidence of such a reactivity. However, this anabolic function of GO is particularly intriguing since amino acid oxidases are usually involved in the catabolic utilisation of their substrates. From this point of view, GO resembles L-aspartate oxidase which converts L-aspartate to iminoaspartate using molecular oxygen or fumarate as electron acceptors, the first reaction in the NAD^+ biosynthesis pathway in bacteria (MATTEVI *ET AL.*, 1999). Furthermore, GO is also the object of particular attention because it can be used in an *in vitro* assay, in parallel to DAAO, to detect and modulate the level of glycine or D-serine in the proximity of NMDA receptors in human brain.

Here we report the crystal structure of *B. subtilis* GO in complex with the inhibitor glycolate at 1.8 Å resolution. Although the inhibitor was found in an unexpected orientation, active site residues that are likely to bind the substrate or to assist in its oxidation have been tentatively identified on the basis of similarities with other related flavoprotein amine oxidoreductases. In fact, the structures of DAAO, MSOX and LAAO have also been resolved (MATTEVI *ET AL.*, 1996; PAWELEK *ET AL.*, 2000; TRICKEY *ET AL.*, 1999; UMHAU *ET AL.*, 2000); it can thus be expected that comparison of their active sites as well as the mode of interaction with the substrate/ligand would provide insights into the similarities and differences in the structure-function relationships of flavoenzymes active on similar compounds. In addition, and with the aim of clarifying the role of GO in *B. subtilis*, the effect of different carbon and nitrogen sources on cell growth and on the level of GO expression has been investigated.

A.3. Experimental Procedures

A.3.1. Growth conditions and preparation of cell extracts

B. subtilis pre-culture was grown aerobically at 37 °C in the dark and under shaking (180 rpm) on a chemically defined, pH-controlled liquid medium (minimal medium) containing 1x minimal salt solution, 0.4% glucose, 0.005% L-tryptophan, 0.2% L-glutamine, 4 mg/mL $FeCl_3$, 0.2 mg/mL $MnSO_4$ and 1% (v/v) trace element salt solution. The 1x minimal salt solution contained 11.5 mM K_2SO_4, 0.8 mM $MgSO_4$, 6.2 mM $K_2HPO_4x5H_2O$ and 3.4 mM sodium citrate, pH 7.0. The 1x trace element salt solution contained 43 µM $CaCl_2$, 12.5 µM $ZnCl_2$, 2.5 µM $CuCl_2x2H_2O$, 2.5 µM $CoCl_2xH_2O$ and 2.5 µM $NaMoO_4x2H_2O$. This pre-culture was then diluted to a final OD_{600} ~ 0.08 in 500 ml of minimal medium (2-liter flasks) and grown for 16 hours as reported above. The cells were collected by centrifugation, suspended in 2 ml of 1x minimal salt solution and finally used to inoculate 600 ml minimal medium supplemented with the appropriate nutrients (initial OD_{600} ~ 0.08). Alternatively, a classic M9 medium (containing 1x minimal salt solution, 22 mM glucose, 2 mM $MgSO_4$ and 0.1 mM $CaCl_2$) supplemented with the appropriate nutrient was also used. The cells were grown in flasks as reported above and collected at different growth phases by centrifugation (4000 rpm for 10 min at room temperature) from 100 ml of fermentation broth. Cell growth was followed by optical density measurements at 600 nm. The crude extracts were prepared according to (Sambrook and Russel, 2001). In detail, 1 gram of cell paste was added to 5 ml of a 2 mg/ml lysozyme solution in TE buffer, pH 8.0, and incubated for 30 min at 37 °C; this sample was then centrifuged at 5000 rpm for 20-30 min at room temperature and 10 ml of lysis buffer added to the pellet (50 mM TrisHCl pH 8.0, 1 mM EDTA, 100 mM NaCl, 1.1 mM PMSF, 5 µg/ml DNase I). After 15-30 min of incubation, the crude extract was recovered by centrifugation at 14500 rpm for 15 min at 4 °C.

A.3.2. Protein analyses, enzyme assay and gel-permeation chromatography

B. subtilis crude extracts were employed for the following assays: a) determination of protein concentration (Biuret method); b) determination of GO activity (polarographic assay, see below); and c) determination of the total GO concentration (by Western blot analysis). A fixed amount of protein (≤ 300 µg) from the crude extract (the soluble fraction obtained after cell disruption and centrifugation) was separated by SDS-PAGE electrophoresis and electroblotted to a nitrocellulose membrane. The same analysis was also performed on whole cell samples by separation of the proteins corresponding to 240 µl of fermentation broth. GO was detected by immunostaining using monospecific rabbit anti-GO antibodies, visualised using both anti-rabbit IgG alkaline phosphatase conjugated with 5-bromo-4-chloro-3-indolyl

phosphate and nitro-blue-tetrazolium chloride as reported in (JOB *ET AL.*, 2002B) or the anti-rabbit IgG horseradish peroxidase conjugate with a chemiluminescent substrate (SuperSignalWest Pico, Pierce Co.). The amount of anti-GO immunoreactive protein was determined by densitometric analysis using a 50- to 1000-ng titration curve obtained from a pure GO.

Glycine oxidase activity was assayed in a thermostated Hansatech oxygen electrode measuring the oxygen consumption at pH 8.5 and 25 °C with 10 mM sarcosine as substrate (JOB *ET AL.*, 2002A; JOB *ET AL.*, 2002B). One unit of GO is defined as the amount of enzyme that converts one µmole of substrate (sarcosine or oxygen) per min at 25 °C.

The oligomerisation state of GO was investigated by means of gel-permeation chromatography on a Superdex™ 200 HR 10/30 column (Amersham) using an elution buffer of 50 mM sodium pyrophosphate, pH 8.5, 5% glycerol and 250 mM NaCl (JOB *ET AL.*, 2002A). The pH effect on GO oligomerisation state was determined chromatographic separation using: 50 mM potassium phosphate (pH 6.5, 7.5), or 50 mM sodium pyrophosphate (pH 8.5), or 25 mM sodium pyrophosphate and 25 mM sodium carbonate (pH 9.5), all containing 250 mM sodium chloride and 5% (v/v) glycerol. These buffers were adjusted to the appropriate pH with HCl or NaOH.

A.3.3. Limited proteolysis experiments

GO (1 mg/ml) was incubated at 25 °C in 50 mM sodium pyrophosphate, pH 8.5, and 1% glycerol with 10% (w/w) trypsin, chymotrypsin or SV8 protease. For electrophoretic analysis, 1.5 mM PMSF was added to protein samples (10 µg of GO) taken at different times after the addition of protease and immediately frozen for analysis by native PAGE on a 7.5% (w/v) polyacrylamide gel or diluted in the sample buffer for SDS-PAGE, boiled for 3 min and then loaded on a 12% (w/v) polyacrylamide gel. Gels were stained with Coomassie Blue R-250 and, only for gels from native PAGE, stained for GO activity as reported in (JOB *ET AL.*, 2002A). The oligomerisation state of proteolysed GO samples was determined by gel-permeation chromatography (see above) and their N-terminal sequence was determined by means of automated Edman degradation using a Procise protein sequencer (Applied Biosystems).

A.3.4. Preparation of the protein and crystallisation

Wild-type GO was expressed in *E. coli* using the pT7-HisGO expression system in BL21(DE3)pLysS *E.coli* cells (JOB *ET AL.*, 2002B). The pT7-HisGO encodes a fully active fusion protein with 13 additional residues at the N-terminus of GO (MHHHHHHMARIRA). The purified protein was concentrated up to 15 mg/ml and equilibrated in 50 mM disodium

pyrophosphate buffer pH 8.5, 10% glycerol by gel permeation chromatography on a Sephadex G-25 (PD10) column.

The recombinant form of GO was crystallised after dynamic light scattering analysis (DynaPro, Protein Solutions Ltd. UK) by the hanging drop vapour diffusion method, mixing 1 µl reservoir and 1 µl protein solution at 18 °C. Crystals with two different space groups were obtained. GO crystals of the space group $P6_122$ were obtained using a 13 mg/ml GO solution and with a reservoir solution containing 1 M sodium citrate (pH 6.2): the hexagonal crystals grew in 2-3 weeks. Crystals of the space group $C222_1$ were obtained using a 13 mg/ml GO solution containing 30 mM sodium glycolate with a reservoir solution containing 100 mM imidazole (pH 8.2), 200 mM calcium acetate and 10 % (w/v) PEG 1000: the orthorhombic crystals grew in 1-3 days. Prior to flash freezing using liquid nitrogen, the crystals were soaked with the corresponding reservoir buffer containing 13% (v/v) ethylene glycol. Heavy atom derivatives were prepared by addition in the reservoir buffer of $KAu(CN)_2$ or $K_2Pt(CN)_4$ to the hexagonal crystal form of GO (1 mM final concentration) and incubating them for 12-16 hours. For cryoprotection, the cryoprotectant buffer, containing 1 mM of the corresponding heavy atom salt was used.

A.3.5. Solution of the GO crystal structure

All datasets were collected under cryogenic conditions at 100 K. For collecting data of the heavy atom derivatives the MAR345 image plate system with a rotating anode x-ray source (Schneider, Offenburg, Germany) was used. A native dataset from the hexagonal crystal was measured at the DESY (EMBL, Hamburg) using the MAR345 image plate system. The 1.8 Å GO-glycolate complex was measured from the orthorhombic crystal at the PSI/SLS (Villigen, Switzerland) using a Mar CCD detector 165 mm in diameter. Space group determination and data reduction was carried out with XDS (KABSCH, 1993). The programs SOLVE/RESOLVE (TERWILLIGER, 2000; TERWILLIGER AND BERENDZEN, 1999) were used to solve phases by multiple isomorphous replacement. With MOLREP (VAGIN AND TEPLYAKOV, 1997), which is part of the CCP4 package (COLLABORATIVE COMPUTATIONAL PROJECT NUMBER 4, 1994), and monomeric sarcosine oxidase as a model (pdb 1el5), the position of one of the two molecules per asymmetric unit was found in the map calculated by RESOLVE. The second molecule was positioned using O (KLEYWEGT *ET AL.*, 2001). Alternating refinement in CNS (BRÜNGER *ET AL.*, 1998) and model building in O was carried out until an R_{free} of around 30% was reached. This model was used with MOLREP to find the four molecules in the asymmetric unit of the orthorhombic crystal. The GO-glycolate complex was refined with REFMAC5 (MURSHUDOV *ET AL.*, 1997) and the secondary structure was analysed with DSSP (KABSCH AND SANDER, 1983). For detailed data collection statistics see Table A.1. and for refinement statistics Table A.2.

Structure plots were produced with the programs MOLSCRIPT (KRAULIS, 1991), RASTER3D (MERRITT AND BACON, 1997) and DINO (PHILLPPSEN A, 2001).

A.3.6. Comparison of GR family members

A superposition for several members of the GR_2 family with GO was performed using the program SUPERIMPOSE (DIEDERICHS, 1995) and parameters describing the superposition were extracted with LSQMAN (KLEYWEGT AND JONES, 1997) from the best topological superposition.

A.3.7. Accession numbers

The coordinates and structure factors of glycine oxidase in complex with the inhibitor glycolate have been deposited in the RCSB Protein Data Bank under accession code 1RYI.Results

A.4. Results

A.4.1. Description of the structure

The GO protein used for the present investigations is a chimeric protein containing 13 residues at the N-terminus in addition to the 369 amino acids in the native form (JOB *ET AL.*, 2002A; JOB *ET AL.*, 2002B). The structure of the complex obtained in the presence of glycolate at 1.8 Å is depicted in Figure A.2 (see also Tables A.1 and A.2). A slightly different protein architecture topology of GO has been previously shown using the 2.3 Å resolution structure (SETTEMBRE *ET AL.*, 2003): the secondary structure topology consists of 14 helices (three small 3/10 helices and eleven regular α-helices) and 18 β-strands (Figure A.4). The position of the residues, that initiate and terminate the secondary structure elements is frequently different in our structure compared with the previously shown topology (SETTEMBRE *ET AL.*, 2003). The main difference is represented by the three newly identified 3/10 helices (depicted in yellow in Figure A.4); the overall topology is not changed. GO is a two-domain protein, which consists of a FAD-binding domain and a substrate-binding domain. The main structural elements are central, antiparallel β-sheets, as first observed in the flavoprotein p-hydroxy-benzoate hydroxylase (SCHREUDER *ET AL.*, 1989). The classic FAD-binding domain is common to the glutathione reductase (GR) class of flavoproteins (KARPLUS AND SCHULZ, 1987). In GO, this motif consists of a six-stranded β-sheet composed of 5 parallel β-strands (strands 6, 2, 1, 10, 18) and one additional antiparallel strand (strand 17) and flanked on one side by three α-helices (helices 1, 7, 10) and on the other side by a three-stranded antiparallel β–sheet (strands 7, 8, 9) and a small alpha helix (helix 8). The polypeptide chain crosses between the two

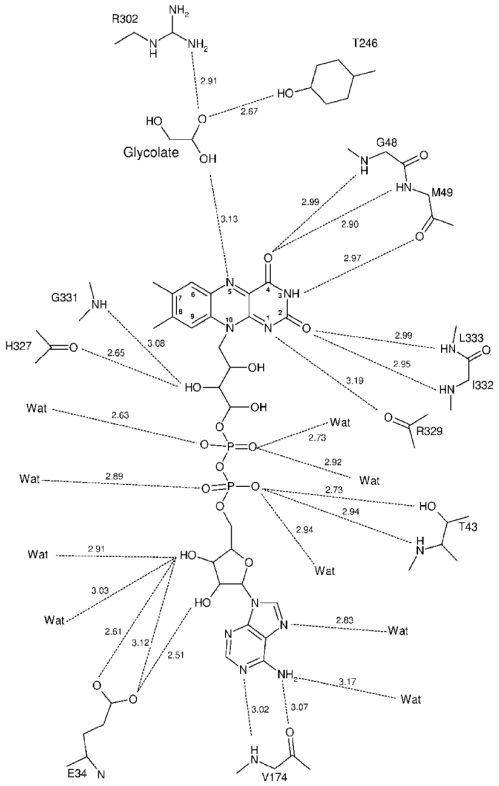


Figure A.1: Schematic representation of the flavin-apoprotein interactions in GO. *View on the si-face of the flavin. Residues interacting with the cofactor via hydrogen bond are depicted. The hydrogen bonds are marked as dotted lines (distances in Å).*

domains four times (after helix 4, strand 5, helix 8 and strand 16). The most significant differences between the GO structure and both MSOX and RgDAAO are represented by α-helix 8, which is missing in the other two enzymes, a different α-helix 3 and 4 topology, which is fused to a single continuous helix in RgDAAO and MSOX, and by the three stranded β–sheet (strands 7, 8, 9 in GO) of the flavin-binding domain, which is conserved in all GR family members and is not conserved in RgDAAO (MATTEVI *ET AL.*, 1996; UMHAU *ET AL.*, 2000) (see Figure A.2). Another main topological difference with RgDAAO is the absence of the loop consisting of 21 amino acids connecting βF5 and βF6 in RgDAAO (Figure A.2b), which is involved in monomer-monomer interaction and is not present in other known DAAO sequences (UMHAU *ET AL.*, 2000). Concerning the catalytic domain of GO an element

corresponding to α-helix 6 of GO is not conserved in RgDAAO. Together with the yeast enzyme, GO shares an active site loop that is shorter by 5-8 residues connecting strands β13 and β14 as compared to pkDAAO and MSOX.

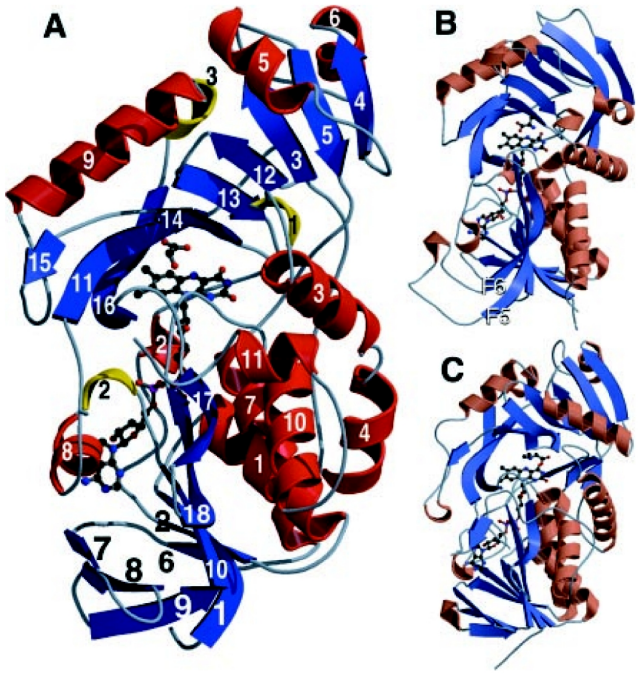


Figure A.2: Ribbon representation of the GO-glycolate complex (1ryi_a) (A), RgDAAO in complex with D-alanine (B), and MSOX in complex with dimethylglycine (C). *Secondary structure elements are highlighted as follows: β-sheets (blue), α-helices (red) and 3/10 helices (yellow).*

At the N-terminus none of the 13 additional amino acids (MHHHHHHMARIRA) present in recombinant GO can be modelled into the electron density, thus apparently possessing a flexible conformation. At the C-terminus, five residues (Glu^{365}, Ala^{366}, Val^{367}, Gln^{368}, Ile^{369}) protrude out of the protein and are not visible in our model and thus do not appear to interact with any of the other subunits. Most interesting, all GO regions that are involved in monomer-monomer interaction (see below) have low temperature factors. In the loop connecting β7 and β8 of the β-meander (Figure A.1) the electron density for four amino acids (Arg^{180}-Ala^{183}) is weak, indicating that part of the loop is very flexible. The region Ala^{55}-Asp^{60} after helix 2 also shows weak electron density. In MSOX this region corresponds to a flexible loop (Tyr^{55}-Tyr^{61}) that changes from a disordered to a weak electron density following the binding of an active site ligand, thus shielding the positive surface potential at the FAD site (Trickey *et al.*,

1999). This loop, and in particular the side chain of Arg[52], was thus suggested to act in MSOX as a switch for active site accessibility (see below).

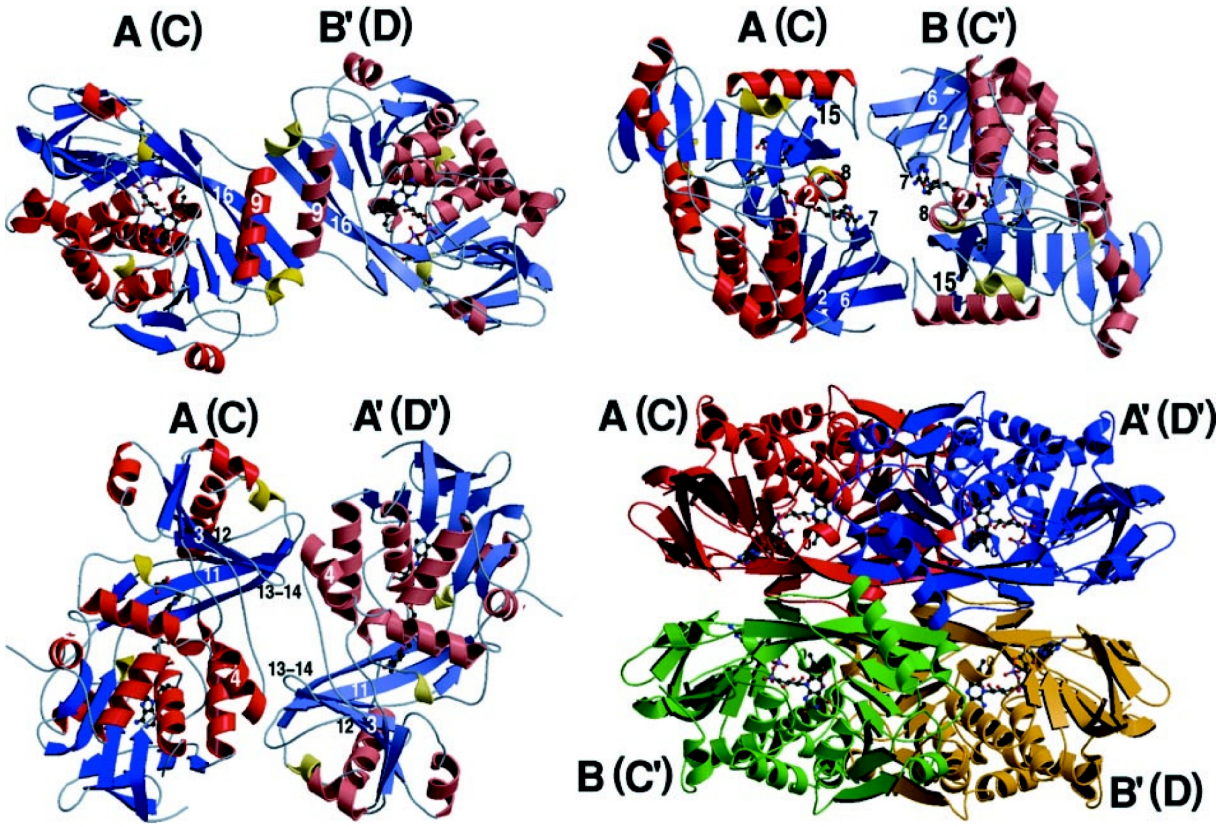


Figure A.3: Different modes of monomer-monomer interaction of the four subunits of tetrameric GO. *Relevant regions are numbered. A: Chain A of the four molecules of the ASU. A': crystallographic symmetry-related molecule of A. B: Chain B of the four molecules of the ASU. B': crystallographic symmetry-related molecule of B. Not shown: the monomers C and D of the ASU make an additional tetramer which their crystallographic symmetry-related counterparts. For details see description in the text. Right lower corner: the two crystallographic tetramers are shown together with the chain identifiers. The crystallographic axis of the tetramer ABA'B' lies vertically and that of the tetramer CDC'D' lies horizontally within the paper plane.*

Table A.1

Summary of data collection, data reduction statistics, and phasing statistics

	GO-glycolate	GO-native	GO-Au	GO-Pt
	PSI/SLS/PX	DESY/BW7B	Rotating anode	Rotating anode
Space group	$C222_1$	$P6_122$	$P6_122$	$P6_122$
Wavelength (Å)	0.97934	0.8463	1.54179	1.54179
Cell dimensions (Å) (a,b,c)	a=73.71 b=218.76 c=217.80	a=b=139.32, c=215.74	a=b=140.20 c=215.88	a=b=140.18 c=215.06
Resolution (Å)	1.8	3.0	3.1	3.5
No. of reflections	581432	153016	327517	188451
No. of unique refl.	159578	25398	42479	29476
Redundancy	3.6	6.0	7.7	6.4
Completeness[a]	98.2% (96.5%)	99.3% (99.8%)	99.6% (100%)	99.4% (99.8%)
R_{sym}[a,b]	4.6% (28.9%)	7.3% (48.4%)	11.2% (43.4%)	19.1% (38.8%)
$R_{merged\text{-}f}$[a,c]	7.1% (34.7%)	7.4% (32.2%)	8.7% (23.4%)	12.9% (22.5%)
I/σ[a]	15.85 (3.67)	16.6 (4.79)	16.71 (5.10)	10.16 (4.83)
Number of heavy atom sites found by SOLVE			5	3
Phasing power of acentric (centric) reflections			0.65 (0.61)	0.82 (0.69)
Figure of merit (FOM)			0.36	
FOM of acentric (centric) reflections after density modification			0.62 (0.66)	

[a] *Values for the outer resolution shell are given in parentheses. Paul Scherer Institute / Swiss Light Source/Protein Crystallography beamline (PSI/SLS/PX), Deutsches Elektronen Synchrotron / Beamline BW7B (DESY/BW7B).* [b] $R_{sym} = \Sigma\Sigma_i |I_i - \langle I \rangle| / \Sigma \langle I \rangle$, *where* $\langle I \rangle$ *is the mean intensity of N reflections with intensities* I_i *and common indices h,k,l.* [c] *(*Diederichs and Karplus, 1997*).*

At the N-terminus none of the 13 additional amino acids (MHHHHHHMARIRA) present in recombinant GO can be modelled into the electron density, thus apparently possessing a flexible conformation. At the C-terminus, five residues (Glu^{365}, Ala^{366}, Val^{367}, Gln^{368}, Ile^{369}) protrude out of the protein and are not visible in our model and thus do not appear to interact with any of the other subunits. Most interesting, all GO regions that are involved in monomer-monomer interaction (see below) have low temperature factors. In the loop connecting β7 and β8 of the β-meander (Figure A.1) the electron density for four amino acids (Arg^{180}-Ala^{183}) is weak, indicating that part of the loop is very flexible. The region Ala^{55}-Asp^{60} after helix 2 also

shows weak electron density. In MSOX this region corresponds to a flexible loop (Tyr[55]-Tyr[61]) that changes from a disordered to a weak electron density following the binding of an active site ligand, thus shielding the positive surface potential at the FAD site (TRICKEY *ET AL.*, 1999). This loop, and in particular the side chain of Arg[52], was thus suggested to act in MSOX as a switch for active site accessibility (see below).

A.4.2. Homology of GO with other amine oxidoreductases

GO exhibits the highest sequence conservation with the β-subunit of heterotetrameric sarcosine oxidase, sarcosine dehydrogenase and dimethylglycine dehydrogenase (24-27% identity), less similarity with the sequences of MSOX and pipecolate oxidase (~ 21% identity), and a modest similarity with DAAO and D-aspartate oxidase (18.4% identity) (JOB *ET AL.*, 2002B). Furthermore, the primary sequence of GO shows a high degree of conservation with the product of *thiO* gene of *Rhizobium etli* (23.0% of sequence identity). *thiO* is the second ORF of four genes (*thiC*, *thiO*, *thiG* and *thiE*) located on plasmid pb which are

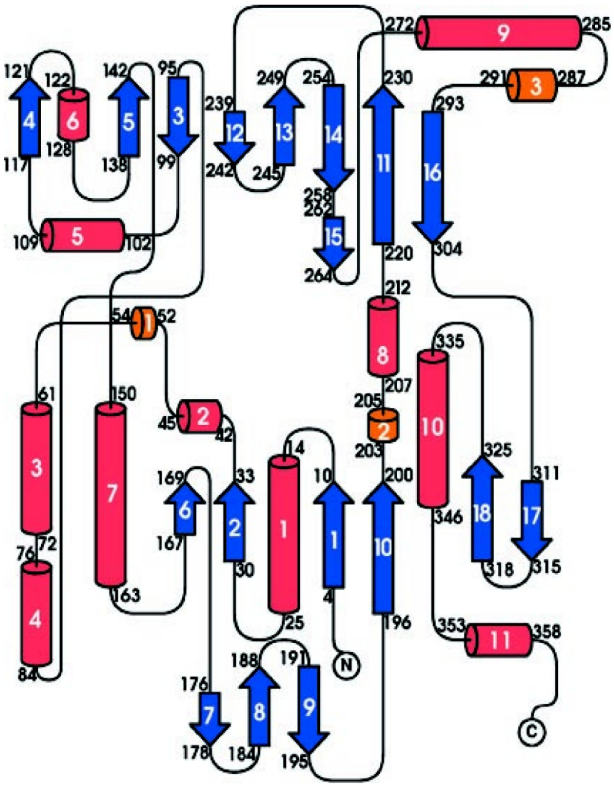


Figure A.4: Secondary structure topology of GO. *α-Helices are shown in red, β-sheets are shown in blue and the newly identified 3/10 helices are shown in yellow.*

involved in the synthesis of thiamine in *R. etli* (MIRANDA-RIOS *ET AL.*, 1997). *R. etli* ThiO protein (a 327 amino acids protein) contains at its N-terminus a flavin adenine dinucleotide binding motif and shares many of the residues involved in the catalytic site of DAAO; it has been also suggested that ThiO may have amino acid oxidase activity (MIRANDA-RIOS *ET AL.*, 1997). In *E. coli* five genes (*thiC, thiE, thiF, thiG* and *thiH*), proposed to be a single transcription unit, are involved in thiamine biosynthesis (VAN DER HORN *ET AL.*, 1993). The ThiO protein of *E. coli* shows a limited sequence identity (12.5%) with *B. subtilis* GO.

By structural overlay, the GO structure was compared to that of other flavoprotein oxidases (see Table A.3). Based on structural and sequence homologies, GO can be classified as a member of the large glutathione reductase (GR) family (all the family members adopt the Rossmann fold) (DYM AND EISENBERG, 2001; ROSSMANN *ET AL.*, 1974) and further into the subgroup GR_2, which was reported to show sequence similarity mainly within 30 residues in the N-terminal region (DYM AND EISENBERG, 2001). Our superposition procedure identified large, structurally homologous parts of GO with DAAO and MSOX (Table A.3). Although comparison of GO with DMGO shows that 284 of 364 residues lie within the distance cutoff of 3.5 Å, the r.m.s. deviation of those residues is high (1.84 Å), reflecting notable structural differences between GO and DMGO. The enzymes pkDAAO, RgDAAO and MSOX could be superimposed with a smaller r.m.s deviation (1.53-1.58 Å) with 222-256 of the 364 residues of the GO structure. This reflects the structural similarity of GO with both DAAOs and MSOX.

Table A.2

Refinement statistics

	GO-glycolate
Resolution (Å)	20-1.8
Total no. of non-hydrogen atoms	12720
No. of water molecules	1046
No. of ligand atoms (FAD and Glycolate)	232
No. of reflections in working set	151580
No. of reflections in test set	8006
R_{work}[a] (%)	17,7
R_{free}[b] (%)	21,4
rms distance from ideal geometry	
Bonds (Å)	0,011
Angles (degrees)	1,32
Ramachandran plot[c]	

Total no. of residues	1456
Most favoured regions	13221
Additionally allowed regions	129
Generously allowed regions	4
Disallowed regions	2

[a] R factor = $\Sigma_{hkl}||F_{obs}|-k|F_{calc}||/\Sigma_{hkl}|F_{obs}|$, were F_{obs} and F_{calc} are the observed and calculated structure factors. [b] For R_{free}, the sum is extended over a subset of reflections excluded from all stages of refinement. [c] (LASKOWSKI *ET AL.*, 1993).

The other GR_2 family members that were compared with GO had either fewer residues lying within a 3.5 Å cutoff, a high r.m.s deviation or a low sequence identity of the superpositioned residues (see Table A.3), thus reflecting the structural and functional difference of these enzymes with respect to GO.

Table A.3

Superposition of GR_2 family members with glycine oxidase.

Protein	PDB accession code (chain name)	Number of residues within 3.5 Å cutoff	r.m.s. deviation (Å) of residues within 3.5 Å cutoff	Sequence identity (%) of residues within 3.5 Å cutoff
DMGO	1pj5	284	1.83	21.8

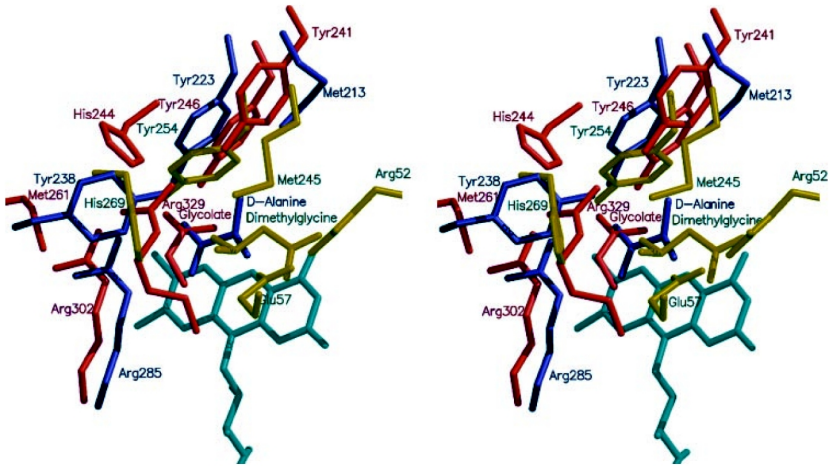


Figure A.5: Stereo picture showing a comparison of ligand-active site interactions (grey, GO FAD). *Blue: RgDAAO in complex with D-alanine. Red: GO in complex with glycolate. Yellow: MSOX in complex with dimethylglycine. FAD cofactors of the three enzymes were superpositioned using LSQMAN.*

MSOX	1el5(a)	256	1.53	23.0

pkDAAO	1aa8(a)	249	1.54	18.4
RgDAAO	1c0p	222	1.58	20.7
PHBH	1bf3	154	1.46	12.3
Phenol hydroxylase	1foh(a)	150	1.70	12.0
Fumarate reductase	1d4d	148	1.75	20.3
D-Aspartate oxidase	1knr	144	1.64	16.7
Glucose oxidase	1cf3	134	1.72	19.4
Cholesterol oxidase	1mxt	128	1.54	17.1
Polyamine oxidase	1b37(a)	114	1.52	17.5
LAAO	1f8r(a)	108	1.41	19.4

A.4.3. FAD binding

Each GO monomer contains one non-covalently bound FAD molecule (JOB *ET AL.*, 2002A; JOB *ET AL.*, 2002B). The FAD-binding patterns of GO, DAAO and MSOX share an overall similarity: in all these enzymes the FAD-binding domain contains the conserved Rossmann fold βαβ motif (β1, α1, β2*)* (ROSSMANN *ET AL.*, 1974), which serves as a dinucleotide-binding motif (WIERENGA *ET AL.*, 1983). The central part of this consensus motif is the sequence GXGXXG (Gly[11], Gly[13] and Gly[16] of helix α1) with the N-terminal end of helix α1 pointing toward the FAD pyrophosphate moiety, as observed for other dinucleotide-binding proteins (WIERENGA *ET AL.*, 1983). The binding of FAD in GO is that typical of the GR_2 subfamily (DYM AND EISENBERG, 2001); the prosthetic group adopts an extended conformation with the isoalloxazine ring located at the interface between the FAD-binding domain and the substrate-binding domain, facing with its *re*-side the inner part of the substrate-binding cavity. The cavity is located distant from the monomer-monomer interaction sites facing towards the bulk solvent. The whole cofactor is buried inside the protein (Figure A.2) and the isoalloxazine ring is not directly solvent accessible. A similar situation was also observed in DAAO (MATTEVI *ET AL.*, 1996; UMHAU *ET AL.*, 2000), while in *p*-hydroxybenzoate hydroxylase the flavin benzene ring is exposed to the bulk solvent, allowing the flavin to adopt two different conformations (ENTSCH AND VAN BERKEL, 1995). The large majority of the potential FAD hydrogen bonds are formed with the protein residues, thus resulting in a tight net as shown in Figure A.1; a K_d value for the apoprotein-FAD complex of $5 \pm 2 \times 10^{-8}$ M has been calculated[§]. The isoalloxazine ring is held in place by a hydrogen bond between its N-3H-C-4=O and the backbone of Gly[48] and Met[49], while N-5 is within hydrogen bond distance to one of the oxygens of the inhibitor glycolate (this flavin position interacts with the backbone NH group of Gly[52] and Ala[49] in RgDAAO and pkDAAO, respectively, and with a water

§ L. Pollegioni, personal communication

molecule in MSOX). The flavin N-1 is within hydrogen bond distance to the Arg^{329} backbone carbonyl oxygen in GO (such an interaction is absent in pkDAAO and corresponds to RgDAAO Ser^{335} and MSOX Lys^{348}). The environment of O-2 position is slightly different between GO and DAAOs. In GO the O-2 forms two H-bonds with the backbone NH group of Ile^{332} and Leu^{333} (Figure A.1). In pkDAAO O-2 interacts with a threonine and with the partial positive charge of a dipole from helix αF5 (MATTEVI *ET AL.*, 1996), and in MSOX the flavin O-2 forms a hydrogen bond to the side chain nitrogen of Lys^{348} (TRICKEY *ET AL.*, 1999).

These interactions serve to stabilize the electrophilic character of the flavin ring, the negative charge of the anionic form of the semiquinone (JOB *ET AL.*, 2002A) and, probably, of the fully reduced flavin of GO. The benzene ring of the isoalloxazine moiety makes van der Waals contacts with a pocket formed by the antiparallel β-strands 11 and 15; the contacting residues are Ala^{45}, Ala^{46}, Ala^{47}, Gly^{224}, Cys^{226}, Tyr^{246}, Ala^{259}, Gly^{300} and Arg^{302}. In particular, Arg^{302} is ~ 4 Å from the C-7 and C-8 position of the flavin. Interestingly, GO Gly^{300} occupies the same position as Cys^{315} in MSOX, the site of covalent flavinylation. The covalent attachment of the flavin cofactor to the apoprotein moiety requires the contribution of a base to act as a proton acceptor from the 8α-CH_3 group of FAD during tautomerisation and possibly to generate the reaction thiolate from Cys^{315} (MATTEVI *ET AL.*, 1996). The proposed candidate is His^{45} in MSOX (in GO an alanine is instead present). Analogously to DAAO and MSOX, there are no acidic residues in the region surrounding the isoalloxazine ring.

The FAD diphosphate group forms one H-bond with Thr^{43} in GO. Five H_2O molecules are found at optimal hydrogen bonding distance with four of the phosphate oxygen atoms (Figure A.1). The side chain of the highly conserved GO Asp^{34} of the FAD-binding domain interacts via three strong hydrogen bonds with the two OH-groups of the AMP ribosyl moiety. The adenine moiety forms hydrogen bonds to the backbone of Val^{174} and a water molecule (Figure A.1). Although different amino acids interact with AMP in GO, MSOX and DAAO, the overall picture is similar.

A.4.4. Mode of oligomerisation

Native GO from *B. subtilis* is a stable 169-kDa homotetramer whose oligomerisation state is not dependent on protein concentration in the 0.01-13 mg/ml concentration range (JOB *ET AL.*, 2002B). The molecular mass of GO in solution estimated from dynamic light scattering and blue native PAGE (SCHÄGGER *ET AL.*, 1994) is ~160 kDa and corresponds to the theoretical value of 173 kDa for the recombinant GO tetramer. Crystals of GO have the space group $P6_122$ and $C222_1$. In both space groups GO crystallised as a tetramer. The tetramers of both space groups are identical and have a 222 point symmetry. The tetramer of the hexagonal crystal is also identical to the one, which was described previously (SETTEMBRE *ET AL.*, 2003). In

the orthorhombic crystal, there are four molecules (monomers) in the asymmetric unit, which build two identical tetramers: the first tetramer consists of the monomers A and B and their symmetry mates A' and B' after applying the symmetry operator (x, -y +2, -z +2) on chain A and B, the second tetramer consists of chains C and D and their symmetry mates C' and D' after applying the symmetry operator (-x, y, -z +3/2). The difference between the two tetramers is the orientation of the non crystallographic axis with respect to the tetramer (see Figure).

In detail, different surface regions are involved in contacts made by each monomer with the other three monomers in the same tetramer. The interface between monomers A and B' (and C and D) is largely due to residues from α9 and β16 of each subunit and buries a total surface of about 1770 Å^2. There are eight hydrogen bonds between the two monomers, formed by the following residues: Pro^{270}-Val^{204}, Leu^{272}-Met^{292}, Gly^{273}-Gln^{290} and Lys^{283}-Glu^{276}. Apart from the last pair of residues, the hydrogen bonds are between main chain N and C atoms. Moreover, the Leu^{275} from α9 forms a hydrophobic pocket with the β16 residues Val^{294} and Phe^{297}. Interestingly, the side chain of Phe^{297} adopts two conformations. Interaction between monomers A and A' (C and D') is due to residues belonging to loop β11-12, loop α4-β3 and loop β13-14 and buries a total surface area of abuout 1870 Å^2; no hydrogen bonds are present, but the loops fit together very tightly. The size of this contact area is around 250 Å^2 larger than reported previously (SETTEMBRE *ET AL.*, 2003). The third interaction between monomers A and B (C and C') is due to residues from loops β15–α9, β2-α2, β6-β7 and Met^{208} from α8, with a total buried surface area of about 975 Å^2. The accessibility from outside to the funnel leading to the active site of each monomer is not restricted by the quaternary structure of GO; the openings face the bulk solvent and are far away from each other. Moreover, the GO quaternary structure shows that interaction between the flavin cofactors of different subunits is not possible. Superposition of the four chains with LSQMAN show, that the r.m.s. deviation among the monomers (A ,B, C, D) is between 0.22 and 0.44 Å. The r.m.s. deviation between both tetramers (ABA'B' and CDC'D') is 0.35 Å. This shows that both tetramers are identical and the r.m.s. deviation is from the observed difference between the intrinsic structure of the monomers.

In order to modify the stable tetrameric structure of GO, different approaches have been used. The tetrameric state of GO is not affected by pH in the pH range 6.5-9.5, as determined by means of gel-permeation chromatography. Treatment with increasing concentrations of thiocyanate (up to 1 M) results in a decrease of the peak corresponding to the tetrameric GO, but a corresponding increase in a peak at ~ 46 kDa corresponding to the monomer was not observed. This result indicates that the lipophilic ion affects the stability of GO in solution but does not alter its monomer-monomer interactions. Tetrameric GO is also strongly resistant to

proteolysis: treatment of GO (1 mg protein/ml) with 10 % (w/w) trypsin, chymotrypsin or SV8 protease (at 25 °C and pH 8.5) for up to 4 hours does not change its elution volume in gel-permeation chromatography. In contrast to GO, the dimeric oligomerisation state of RgDAAO is affected by proteolysis (POLLEGIONI *ET AL.*, 1995). Only the N-terminus of GO is sensitive to proteases: Western blot analysis and protein sequencing demonstrated that trypsin and chymotrypsin treatment, respectively, converted the 382 amino acids of recombinant GO monomer into a protein form that was shorter by 8-10 and 12-14 residues (lacking the His-tag at the N-terminus).

In contrast to this, the apoprotein form of GO is monomeric and rapidly converts to tetrameric holoenzyme upon addition of FAD[§]. Considering that the FAD cofactor is involved in many protein core contacts (see above) and that the conversion to the apoprotein form abolishes the CD signal in the near-UV due to the tertiary structure of the holoenzyme, it was deduced that the organisation of the quaternary structure follows holoenzyme reconstitution.

A.4.5. The active site

The active site of GO is a cavity delimited by the two long β-strands, 11 and 16, bent around the isoalloxazine ring of the flavin and the two short β-strands, 13 and 14, located close to the substrate binding site; the flavin forms the bottom of the cavity (Figure A.5). The electron

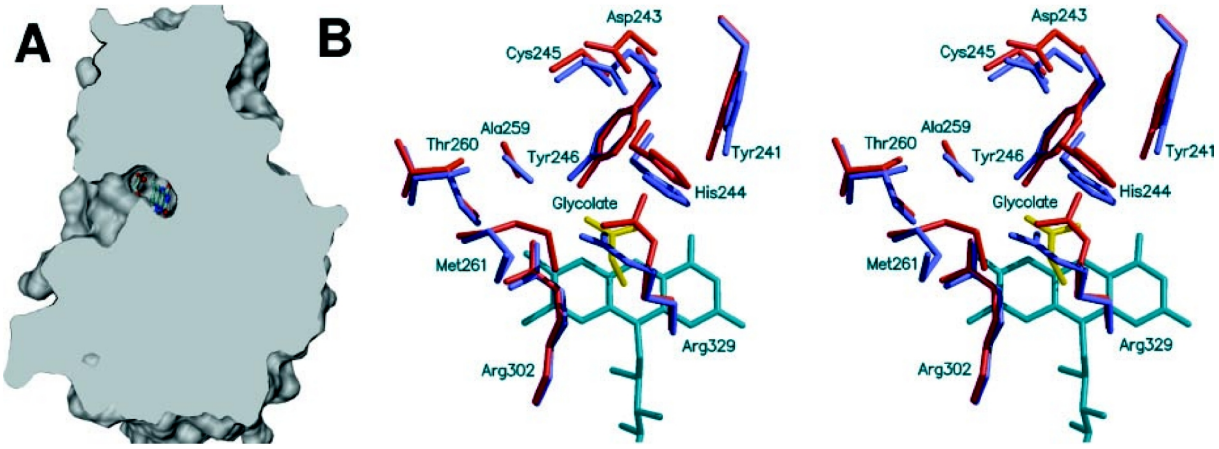


Figure A.6: (A)Side view of the substrate channel to the active site. *Part of the flavin and the inhibitor glycolate are visible within the active site.*

(B). Comparison of the active site entrance between GO in complex with glycolate (1ryi_a, red) and the free form of GO (1ng4, blue) *showing a magnified view along the direction of the funnel leading to the active site. The glycolate is depicted in yellow and the GO FAD in grey. Superposition was performed using SUPERIMPOSE (DIEDERICHS, 1995).*

§ E. Pollegioni, personal communication

density is clearly seen for all atoms of the ligand glycolate in all four monomers of the asymmetric unit. The interactions can be described as follows: a) Arg^{302} and Tyr^{246} form hydrogen bonds with one oxygen of the α–COO$^-$ of the inhibitor glycolate (2.91 and 2.67 Å); b) the second oxygen of the inhibitor α–COO$^-$ group is at hydrogen bond distance to the N-5 of the FAD; c) one terminal nitrogen of Arg^{329} and the main chain carbonyl of His^{244} form a hydrogen bond (3.02 Å) and are in close proximity to the α–COO$^-$ of the bound ligand, probably preventing its rotation. Comparison of the structure of glycolate bound to GO with those of D-alanine in DAAO (UMHAU *ET AL.*, 2000), dimethylglycine in MSOX (TRICKEY *ET AL.*, 1999), and with acetylglycine in GO (SETTEMBRE *ET AL.*, 2003) shows that the α–COO$^-$ of glycolate does not make two hydrogen bonds to the guanidyl group of GO Arg^{302}, but instead that the ligand is unexpectedly rotated by ~120° and binds in a different orientation, which is observed in all four molecules of the ASU. This binding mode may be favoured by the small size of glycolate, and probably could resemble the binding mode of glycine, which has approximately the same size as the inhibitor we used.

It is noteworthy that the substrate binding geometry of MSOX is reversed compared to DAAO and GO (Figure A.5). GO Tyr^{246}, which forms with its side chain oxygen a hydrogen bond to one of the α-COO$^-$ oxygens of glycolate, is found at a position resembling that of RgDAAO Tyr^{223} (a residue involved, mainly by steric effects, in substrate binding) and of MSOX Tyr^{254} (whose role in substrate binding is considered marginal). GO Arg^{302}

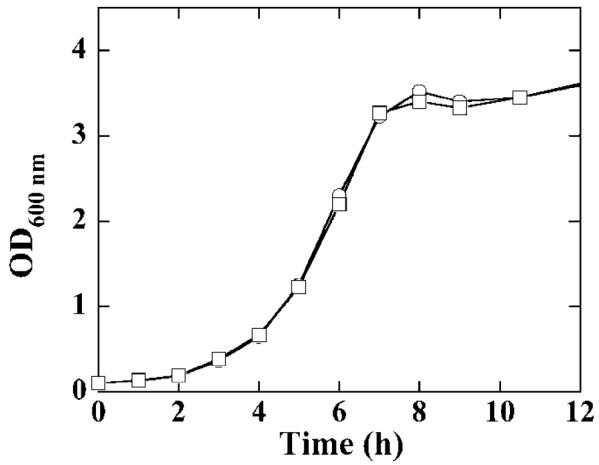


Figure A.7: Effect of thiazole alcohol on the growth of wild-type strain of B. subtilis. *Growth of B. subtilis was performed in minimal medium in the absence (o) and in the presence (□) of 2 μM thiazole alcohol.*

corresponds to RgDAAO Arg^{285} (and pkDAAO Arg^{283}) and to MSOX Arg^{52}: in all these

enzymes, it is the residue that interacts with the α-COO^- of the substrate. In MSOX the movement of Arg^{52} following inhibitor binding induces a large replacement of a loop region which directs MSOX Glu^{57} into the active site (to bind MSOX Arg^{52}) (MATTEVI *ET AL.*, 1996). Concerning the modulation of the substrate specificity, GO Tyr^{241} is located at a position resembling that of MSOX Met^{245} and RgDAAO Met^{213} (Figure A.5). This latter residue was recently demonstrated to modulate the substrate specificity of yeast DAAO (SACCHI *ET AL.*, 2002): the introduction of a positive charge at this position by site-directed mutagenesis provided a DAAO variant active on both neutral and acidic D-amino acids.

Most interesting, the loop found in pkDAAO and which was proposed to act as lid-controlling access to the active site (MATTEVI *ET AL.*, 1996), is absent in GO (as well as in RgDAAO and MSOX). The pkDAAO loop contains an important residue, Tyr^{224}, which is probably involved in substrate/product fixation (MATTEVI *ET AL.*, 1996). The RgDAAO Tyr^{238} side chain, which was also proposed to change its position in order to allow substrate/product exchange, is placed at a position similar to that in pkDAAO Tyr^{223}, GO Arg^{329} and MSOX His^{269} (Figure A.5). The latter residue was preliminary considered the active site base, although now a different role has been proposed (ZHAO AND JORNS, 2002).

No lid is evident in GO, but some side chain rearrangements occur upon inhibitor binding: in particular in Arg^{329} and Met^{261}. We proposed that His^{244} should be involved in a system to drive the protons outside the active site: a proton should be taken up from the substrate by the latter residue and transferred to Arg^{329} and to Met^{261} that is outside the active site. In GO, the following residues form the entrance to the active site: Ala^{259}, Thr^{260}, Met^{261}, Arg^{329}, His^{244}, Asp^{243} and Cys^{245} (Figure A.6). Site-directed mutagenesis of Met^{261} (to Tyr and His) supports such a conclusion. The mutants have properties quite similar to the wild-type (similar V_{max}); the main change is a 10-fold increase in K_m for the substrates and in K_d for the inhibitors, and a change of the flavin redox properties in the free enzyme form§. There is a striking difference between our structure and the conclusions derived by the unliganded structure (SETTEMBRE *ET AL.*, 2003) regarding the seal of the active site. Comparison of free GO with GO in complex with glycolate (Figure A.6) shows that the movements of residues Met^{261} and Arg^{329} do not fully protect the bound ligand from external solvent, and thus the imine product from hydrolysis.

A.4.6. Effect of different carbon and nitrogen sources on B. subtilis growth and on GO expression

A previous report (SETTEMBRE *ET AL.*, 2003) showed that the *B. subtilis* mutant *thioO*$^-$ was unable to grow in a synthetic medium not containing thiazole. The growth curves of the wild-

§ L. Pollegioni, personal communication

type strain of *B. subtilis* on minimal medium in the presence and in the absence of thiazole alcohol are identical (see Figure A.7), thus demonstrating that wild-type *B. subtilis* does not require the thiazole moiety of thiamine to grow. The activity level of GO during the time course of *B. subtilis* growth in the presence and in the absence of thiazole is constant and corresponds to the basal level (0.03 U/gram cell).

We used a chemically defined growth medium to elucidate the effect of nutrients on the level of GO synthesis. *B. subtilis* cells were grown on a minimal medium (that contains 0.2% L-glutamine), as well as on M9 medium (according to (SETTEMBRE *ET AL.*, 2003)), supplemented with different carbon and nitrogen sources (see Table A.4). Using M9 minimal medium, *B. subtilis* grows following the addition of ammonium chloride or D-alanine as a nitrogen source, while no growth is observed when either glycine or sarcosine are the sole nitrogen source. However, growth of *B. subtilis* was observed in the presence of both glucose and glycine using the minimal medium from (SETTEMBRE *ET AL.*, 2003) because of the presence of 0.2% L-glutamine in this chemically defined medium. Glycine cannot be used by *B. subtilis* as the sole carbon source in either M9 or minimal medium, while sarcosine supports *B. subtilis* growth as the carbon source only using the minimal medium (because it contains L-glutamine as the main nitrogen source, see above). On the other hand, the D-isomer of alanine can be efficiently used as a carbon source (even using the M9 medium) in the presence of ammonium chloride as a nitrogen source, and also as a nitrogen source with glucose as a carbon source. It is important to note that, with the exception only of GO, no enzymatic activities have been reported in *B. subtilis* that can efficiently and directly catabolyse D-alanine.

The level of GO activity and total GO protein expressed in *B. subtilis* was investigated for all the conditions reported previously that supported the growth of *B. subtilis* cells. In all cases, the GO activity in the crude extract is at the limit of detection: during all growth phases the amount of GO corresponds to the basal level of enzymatic activity (0.03-0.04 U/liter fermentation broth). Interestingly, Western blot analysis showed that GO in the crude extract was always below the detection limit under these growth conditions (i.e. the GO expression is < 0.03% of the total proteins present in the crude extract). The same result was also obtained when Western blot analysis was performed using the whole cells.

A.4.7. Effect of phosphate and thiamine pyrophosphate on GO activity

A previous, detailed kinetic characterisation of GO showed that the enzyme follows a classic Michaelis-Menten dependence of the reaction rate as a function of substrate concentration, without any indication of sigmoidal behavior (MOLLA *ET AL.*, 2003). In the previously reported structure of GO (SETTEMBRE *ET AL.*, 2003), a phosphate ion was modeled into a buried, positively charged pocket formed by interaction of two monomers in proximity of the

residues Arg[89], ArgH, and Arg[296]. In the orthorhombic crystal all three corresponding sites, which lay between the monomer pairs AA', BB', and CD where investigated, neither those sites, which are located on a crystallographic axis (AA', BB' nor those on a non crystallographic axis (CD) show any indication of bound phosphate ions at these sites. In agreement with our crystallographic data, the addition of phosphate (in the 1-130 mM concentration range) does not affect the activity of GO.

Because of the proposed role in *B. subtilis* of GO in the biosynthesis of the thiazole moiety of thiamine (SETTEMBRE *ET AL.*, 2003) and of its tetrameric oligomerisation state, the effect of the proposed final product (thiamine and thiamine pyrophosphate) of this synthetic pathway on the enzymatic activity of GO was also assessed. As observed previously for phosphate, even the addition of 0.1, 1 and 50 mM thiamine and thiamine pyrophosphate does not affect the activity of GO.

A.5. Discussion

Amine oxidation is widespread in biology, and several different enzymes have evolved that catalyse these reactions. In the present investigation we used a combination of structural and functional studies to clarify the physiological role of GO in *B. subtilis*. On the basis of the data presented here, we propose that the structure of *B. subtilis* GO resembles that observed for MSOX and DAAO and that both enzymes have a catabolic role, although in the specific mode of FAD and substrate binding different amino acids are involved. On the other hand, a different role of GO in the formation of the thiazole moiety of thiamine was recently proposed (SETTEMBRE *ET AL.*, 2003); this role requires that the imine product of dehydrogenation of glycine is trapped at the active site of GO with the thiocarboxylate produced by the following enzyme of the metabolic pathway, ThiS. This proposal is also supported by the sequence homology between GO and ThiO, a protein that has been suggested to catalyse this reaction in *R. etli* (MIRANDA-RIOS *ET AL.*, 1997). This nucleophilic addition may occur before hydrolysis of the imine, which, when it is released from the enzyme, takes place immediately. If such a condensation reaction really occurs at the active site of GO (SETTEMBRE *ET AL.*, 2003), it should be more easily supported by a ping-pong kinetic mechanism of the GO reaction since the rate of product release from the reduced flavin form is known to be slower than that from the re-oxidised enzyme (MOLLA *ET AL.*, 2003). Instead, we recently demonstrated that GO follows a sequential, ternary-complex kinetic mechanism with glycine as well as with sarcosine and D-proline as substrate (MOLLA *ET AL.*, 2003); the kinetic mechanism is consistent with the flavin reoxidation starting from the reduced enzyme-imino acid complex. Furthermore, comparison of the active site entrance of GO in complex with glycolate and in the free form (Figure A.6) does not show any seal of the active site, that would prevent the hydrolysis of the imine

product before release. The entrance of the substrate to the active site of GO is found to be on the solvent-accessible side of each monomer, far away from monomer-monomer contact areas. GO does not have a lid which controls substrate access and product release, a main difference in comparison to that observed in other known amino acid oxidases. The structural data do not clarify how GO can interact with ThiS and how the imine product can react with the thiocarboxylate intermediate bound to ThiS. Only further investigations concerning the structure of the proposed complex between GO and ThiS can provide a rationale for the proposed role of GO in thiamine synthesis if such a complex demonstrated a direct connection between the two active sites, together with a sealing from solvent of GO active site.

A further obscure point that hardly correlates with the proposed role in the anabolic production of thiamine is represented by the (wide) substrate specificity of GO. Is it only an evolutive memory? GO is active on many of the compounds specifically oxidised by DAAO, MSOX and monoamine oxidase (JOB *ET AL.*, 2002A; JOB *ET AL.*, 2002B). In addition, the V_{max} is similar to that of many of the substrates tested, because it is essentially due to the rate of product release from the reoxidised enzyme (MOLLA *ET AL.*, 2003). The rate of flavin reduction by the substrate is similar among glycine and sarcosine (the best substrates of GO), and indeed the K_m for sarcosine is slightly lower than for glycine (MOLLA *ET AL.*, 2003).

An additional intriguing point pertains to the oligomeric state of GO. Why is GO tetrameric, whereas all the other members of the amine flavooxidase family are monomeric or dimeric (no other members of the GR_2 subfamily are tetrameric)? We demonstrated that GO does not show any allosteric behaviour: the enzyme possesses classic Michaelis-Menten kinetics as a function of substrate concentration (MOLLA *ET AL.*, 2003) and no effect on the enzymatic activity by phosphate, thiamine and thiamine pyrophosphate is observed.

In order to clarify the role of GO in *B. subtilis*, the effect of different nitrogen and carbon sources on the growth and the induction of GO activity were also investigated. A wild-type strain of *B. subtilis* does not require thiazole to grow, thus confirming that it is able to synthesise thiamine. We also demonstrated that glycine and sarcosine, two putative *in vivo* GO substrates, do not support *B. subtilis* growth when used as the sole carbon or nitrogen source, while D-alanine can be used as the main carbon source (Table A.4). This latter compound can only be directly used if it is catabolised by GO, since no other enzymatic activities have been reported in *B. subtilis* to deaminate D-amino acids. Otherwise, D-alanine could be converted to the corresponding L-isomer by an amino acid racemase (various racemases have been identified in *B. subtilis* genome, in particular a D-alanine racemase) and thus L-alanine should also be used as a carbon source. This latter possibility is in good agreement with our experimental evidence showing that the amount of GO (activity and protein) is not increased in the absence of thiazole in combination with using D-alanine as the sole carbon source, i.e. it is not inducible. The evolution of the regulation of amino acid

degradative enzymes in *B. subtilis* resulted in enzymes present at high levels in sporulating cells and spores, rather than preferentially expressed during nitrogen-limited growth (for a review see FISHER, 1999). Catabolite-repressed genes in *B. subtilis* are controlled by more than one global regulatory mechanism. Although none of the *Bacillus* catabolite repression mechanisms is understood at the molecular level, it is clear that they are not analogous to the cAMP-CRE-dependent mechanism that is operative in *E. coli* (for a review on catabolite repression and inducer control in Gram-positive bacteria, see SAIER AND RAMSEIER, 1996 AND SAIER *ET AL.*, 1996). We performed a search for a 14-bp palindromic sequence element corresponding to the catabolite-responsive element (consensus sequence: TGT/A*A*CG*T*A/TCA), which mediates catabolite repression in a number of *B. subtilis* genes (HUECK AND HILLEN, 1995). Catabolite responsive elements show a striking variability in their sequence and position with respect to the transcriptional start sites of regulated genes. Two sequences similar to the consensus sequence for catabolite responsive elements are found within the reading frame of the GO-coding gene (close to the 5' start site): with respect to the observed deviations of such elements from the canonical sequence (at least at three positions) GO cannot be considered as a protein active in carbon catabolite pathways.

Table A.4

Effect of carbon and nitrogen sources on the aerobic growth of B. subtilis on chemically defined media: minimal medium (containing 0.2% L-glutamine) and M9 medium (in parentheses), at 37 °C and 180 rpm.

		Carbon source (20 mM)			
		Glucose	Glycine	Sarcosine	D-alanine
Nitrogen source (18 mM)	NH_4Cl	+ (+)	(-)	(-)	(+)
	Glycine	+ (-)	-		
	Sarcosine	(-)			
	D-Alanine	+ (+)			
	L-Glutamine			+	

We also showed that glycine and sarcosine cannot be used as a nitrogen source, while D-alanine supports *B. subtilis* growth even in the absence of L-glutamine (Table A.4). Despite the different ecological habitats and life cycles of *B. subtilis* and enteric bacteria, such as *E. coli*, only minor differences in the physiology of ammonium assimilation have been reported between these bacteria (SCHREIER, 1993). *B. subtilis* has no assimilatory glutamate dehydrogenase activity; therefore, ammonium assimilation occurs solely by the glutamate synthase pathway (its synthesis is raised during nitrogen-limited growth). Glutamine is the

best nitrogen source for *B. subtilis*, followed by arginine, and even ammonium is a good nitrogen source. Expression of the arginine, proline and histidine degradative enzymes has been reported to be substrate inducible in both *B. subtilis* and enteric bacteria (FISHER, 1993). However, in contrast to the case for *E. coli*, their expression is not subjected to nitrogen regulation in *B. subtilis*, raising the possibility that this bacterium does not contain any global nitrogen-regulatory system. Our results show that the expression of GO is not modified by the different nitrogen sources used and in particular not by glycine, sarcosine and D-alanine even if this latter compound can used as the main nitrogen source and allows *B. subtilis* growth. Taken together, our microbiology experiments exclude regulation of GO synthesis according to the composition of the medium in *B. subtilis*, and therefore do not support a main role of GO in the catabolic utilisation of primary and secondary amines.

In conclusion, the structural and functional properties determined on GO demonstrate that it belongs to the amino acid oxidase class of flavoproteins (in particular to the GR_2 subfamily), and that it is characterised by a broad substrate specificity, low kinetic efficiency and a unique and stable tetrameric oligomerisation state (depending on FAD binding). The combination of these investigations does not clarify how GO is involved in thiamine biosynthesis (mechanistic limitations of this role have been also reported by SETTEMBRE *ET AL.*, 2003), nor do the microbiology experiments support a main catabolic role for this flavooxidase in *B. subtilis*.

B. Expression and refolding experiments of neuronal cell adhesion molecules

This work resulted in soluble TAG-1; the crystal structure is presented in the chapter C.

B.1. Introduction

B.1.1. Axonal cell adhesion molecules

The development of the nervous system and its function is dependent on the formation of a huge number of connections, which link spatially separated cells to a complex network of neurons. To reach distant targets, the axons grow along predetermined pathways, which are

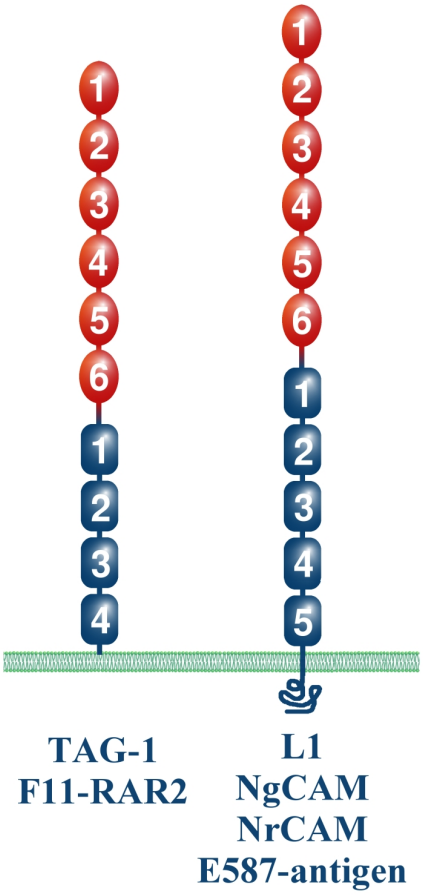


Figure B.1: Schematic representation of the domain composition of the F11 and L1 family. *The F11 family members are composed of six immunoglobulin-like domains and four fibronectin type III domains, attached C-terminally to the membrane via a GPI anchor. The L1 family exhibits a fifth fibronectin type III domain, one transmembrane helix, and a small intracellular domain.*

defined by signal molecules that act as so called guidance cues, a term which describes proteins, that transfer guidance information to growing axons. These guidance cues act as attractive and repulsive components for axons on their way to their targets. An important class

of molecules that act as guidance cues on one hand, and as receptors for those cues on the other hand, are the cell adhesion molecules of the immunoglobulin (Ig) superfamily (DODD AND JESSELL, 1988A; HYNES AND LANDER, 1992). They consist of Ig-like domains and fibronectin-type-III-like (FnIII-like) domains and act mainly on the extracellular side of the cell membrane. The F11 family, a subfamily of the Ig superfamily, contain its founder F11, which is also known as F3 or contactin (RATHJEN *ET AL.*, 1987), and TAG-1 (DODD *ET AL.*, 1988), which is known under different names, depending on authors and species origin. TAG-1 is known in chicken (*Gallus gallus*) as Axonin-1 (ZUELLIG *ET AL.*, 1992) and TAG-1 in human (TSIOTRA *ET AL.*, 1993). All F11 family members have six N-terminal immunoglobulin-like domains in common and four fibronectin-III-like domains, and they are attached to the cell membrane by a glycosylphosphatidylinositol anchor (see Figure B.1, page 32). The L1 family contains human L1, which is also known as L1CAM, CD171 (MOOS *ET AL.*, 1988). Other members are the chicken homologue NgCAM (**N**euron **g**lia **c**ell **a**dhesion **m**olecule) (RATHJEN *ET AL.*, 1987), the goldfish homologue E587-antigen (GIORDANO *ET AL.*, 1997), the fruit fly homologue neuroglian (BIEBER *ET AL.*, 1989), and chicken NrCAM (**N**euron glia cell adhesion molecule **r**elated **c**ell **a**dhesion **m**olecule) (BURGOON *ET AL.*, 1991) which shares 37% sequence identity with NgCAM. The L1 subfamily members are composed of SIX N-TERMINAL IG-LIKE DOMAINS, AND FIVE FNIII-LIKE DOMAINS. ITS MEMBERS ARE MEMBRANE PROTEINS WITH ONE TRANSMEMBRANE HELIX AND A SMALL, CONSERVED INTRACELLULAR PART.

B.1.2. Chicken Axonin-1, NgCAM, and NrCAM guide the axons to the contralateral site

Higher organisms have a bilaterally symmetric central nervous system. For information transfer between both sides it is necessary, that at least some connections exist between them. At an early stage during brain development such connections are made by commissural ganglia which establish projections on the contralateral side. For neurons and their axons it is important to find the correct pathway. Commisural ganglions which grow towards the floor plate express Axonin-1 on their growth cones. For effective floor plate passage of the neurites, an interaction of Axonin-1 of the commissural ganglions with NrCAM expressed by the cells of the floor plate is essential. This was shown by experiments, where either soluble Axonin-1, NrCAM specific, or Axonin-1 specific antibodies reduced the amount of axons, which crossed the floor plate (LUSTIG *ET AL.*, 1999; STOECKLI AND LANDMESSER, 1995). Those axons which did not cross the floor plate, grow rostrally along the midline on the ipsilateral side. Interestingly, NgCAM specific antibodies used in a similar experiment, did not lead to a reduced amount of crossing axons, but instead, the axons lose the ability to stay attached to each other, a phenomenon which was called defasciculation (STOECKLI AND LANDMESSER, 1995). Axonal entering of the floor plate requires an interaction between NrCAM and TAG-1 which

has been shown by experiments with cultured spinal chord explants of commisural ganglia/growth cones and floor plate cells: the interferance of chicken TAG-1 or NrCAM specific antibodies and of soluble chicken TAG-1, resulted in loss of axon growth into the floor plate cells and, as an additional effect of TAG-1 specific antibodies, a disturbed growth cone morphology was observed (STOECKLI *ET AL.*, 1997). In vitro, it was shown, that outgrowing axons do not prefer NgCAM to NrCAM or vice versa, but that a combined NrCAM/NgCAM substratum is preferred to a simple NgCAM or NrCAM substrate (FITZLI *ET AL.*, 2000).

B.1.3. Differences between chicken and rodent expression pattern

A similar function of NrCAM is observed in the rodent developing nervous system, but an important difference in chick and rodent developing nervous system is the expression pattern of NgCAM/L1 and TAG-1. Whereas in chick, both molecules are expressed by commissural ganglia on both sides of the floor plate, namely before entering the floor plate and after its crossing, rodent L1 expression was found only in axons, that had contacted the floor plate. Those L1 expressing axons had crossed the floor plate and projected on the contralateral side, or just contacted the floor plate, but turned rostrally on the ipsilateral side (DODD *ET AL.*, 1988). Rodent TAG-1 expression however was found only in those commissural axons, that had no contact to floor plate cells, these results suggest, that TAG-1 expression is down regulated after floor plate contact (DODD *ET AL.*, 1988).

B.1.4. L1 cell adhesion molecule

THE HOMOPHILIC AND HETEROPHILIC BINDING CAPABILITY OF L1 IS THE BASIS OF ITS VERSATILE ROLE WITHIN THE LIGAND-RECEPTOR NETWORK OF GUIDANCE FORCES THAT ARE INVOLVED IN AXONAL PATH FINDING, GROWTH, AND GUIDANCE. THEREFORE IT IS NOT ONLY A MERE CELL ADHESION MOLECULE, BUT SHOULD BE CONSIDERED AS A NEURONAL RECOGNITION MOLECULE (KENWRICK *ET AL.*, 2000).

L1 IS INVOLVED IN AXON MYELINATION, AS ON ONE HAND NON-MYELINATING SCHWANN-CELLS AND UNMYELINATED AXONS HAVE L1 PRESENT ON THEIR SURFACE, WHICH DISAPPEARED UPON SUCCEREDING MYELINATION (MARTINI AND SCHACHNER, 1986). ON THE OTHER HAND, L1 SPECIFIC ANTIBODIES WERE ABLE TO DISTURB AXON MYELINATION IN AN EXPERIMENT WITH CO-CULTURED SCHWANN CELLS AND DORSAL ROOT GANGLIA (WOOD *ET AL.*, 1990). THE MOLECULAR BASIS OF L1 FUNCTIONALITY IN MYELINATION OF AXONS HAS BEEN REPORTED TO BE DEPENDENT ON AXONAL L1 INTERACTION WITH INTEGRINS LOCATED ON SCHWANN-CELLS (ITOH *ET AL.*, 2005).

Human L1 is involved in an inheritable genetic disease, called the L1-disease, which results in phenotypes like X-linked hydrocephalus, mental retardation, aphasia, shuffling gait, adducted thumbs (MASA-syndrome), or spastic paraplegia type 1 (SPG1) (for a review see (FRANSEN *ET AL.*, 1997)). Affected patients suffer from mental retardation, flexion deformations of the

thumbs, and symptoms of brain malfunction like hydrocephalus, and defects in the corticospinal tract and the corpus callosum (BRÜMMENDORF *ET AL.*, 1998). Many mutations of the L1 gene were identified, which result in these developmental malformations. It was shown in two *in vitro* studies, that L1 heterophilic interaction with TAG-1 as well as its level of expression is affected by those mutations (DE ANGELIS *ET AL.*, 1999; DE ANGELIS *ET AL.*, 2002). Recently, new pathological missense mutations have been identified (SIMONATI *ET AL.*, 2006): the investigation of a boy suffering from L1-disease, resulted in the identification of a mutation of the L1 gene, located within the second FnIII domain. From the fact that L1 mutations affect heterophilic TAG-1 interaction (DE ANGELIS *ET AL.*, 2002), it is speculative but possible to derive that other L1 signal and receptor functions may be affected too. The manifold processes during development, which are identified to be dependent on correct L1 function and its presence give a coherent picture of the importance of L1. Beside the influence of missense mutations on L1 function and interaction, pharmacological doses of ethanol also influence L1 mediated cell adhesion (WILKEMEYER *ET AL.*, 2003) which can cause the so called fetal alcohol syndrome in children of heavy drinking mothers during pregnancy, the most important cause of mental retardation (ABEL AND SOKOL, 1991). In the same field, it was shown that ethanol has an inhibitive effect on L1 mediated signaling of nitrogen-activated kinases namely ERK1/2 mediated phosphorylation of FGFR1 (TANG *ET AL.*, 2006).

B.1.5. Chicken F11

The neuronal cell adhesion molecule F11 from chicken (RATHJEN *ET AL.*, 1987) is a close homologue to Axonin-1 the chicken ortholog of TAG-1 (51% sequence identity) and is also known as F3 and contactin. It is involved in neurite fasciculation and supports axon aggregation (RATHJEN *ET AL.*, 1987). In contrast to chicken TAG-1, only a very small fraction is found in soluble form, whereas the majority is anchored to the membrane (MOSS AND WHITE, 1992). F11 is found in a proprioceptive subpopulation of dorsal root ganglia interacting with NrCAM. In contrast, chicken TAG-1 is found on nociceptive ganglia interacting with NgCAM (PERRIN *ET AL.*, 2001). Although TAG-1 and F11 are capable of interacting with NrCAM and NgCAM directly, both molecules interact *in vivo* specifically with either NrCAM or NgCAM. This selective expression and interaction pattern is important for the development of the projections of sensory axons (FALK *ET AL.*, 2002).

It was shown that F3/contactin acts as a functional ligand of Notch via a *trans* interaction mechanism, which triggers γ-secretase-dependent nuclear translocation of the Notch intracellular domain (HU *ET AL.*, 2003). Notch is a transmembrane receptor mediating cell-cell communication and is responsible in determining cell fates and pattern formation (for reviews see (LAI, 2004; MIELE, 2006) AND A MINI REVIEW ABOUT ITS F3/CONTACTIN INTERACTION (HU *ET AL.*,

2006)). Interaction of axonal F3 with Notch located on oligodendrocytes is proposed to regulate myelination as a positive instructor (HU *ET AL.*, 2003).

B.1.6. Chicken and human TAG-1

The structure of human TAG-1 was solved and a paper (see chapter C), proposing an alternative model for homophilic interaction is in submission. The introduction of chapter C gives a detailed overview of chicken and human TAG-1.

B.1.7. Immunoglobulin domains, properties and considerations

Neuronal Cell adhesion molecules of the L1 and F11 family are composed of Ig and FnIII domains. The extracellular part contain four to five C-terminal FnIII and six N-terminal Ig domains. The Ig domain topology was compared with the Greek key motif, because the arrangement of the secondary structure elements (β-strands as part of two β-sheets) have distinct similarity with patterns found on ancient Greek pottery (RICHARDSON, 1981) (see Figure B.3). The common feature of Ig domains is a sandwich like arrangement of its two β-plated-sheets, which are composed of seven to nine β-strands (LESK AND CHOTHIA, 1982). Although the β-sandwich is a common feature of Ig domains, only four of the β-strands form a common core, whereas the other strands are oriented much more variably. In many cases, the β-sandwich is stabilized by one or more disulfide bridges (first domain of growth hormone receptor, see BAZAN, 1990), and the cysteine carrying strands can vary and do not have to be part of the common core. For domain 1 of CD2 no disulfide bond was reported (JONES *ET AL.*, 1992; examples reviewed by BORK *ET AL.*, 1994).

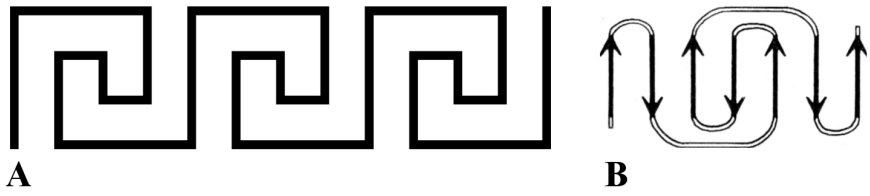


Figure B.2: The Greek key motif. *(A) Schematic picture of the Greek key as found on ancient pottery. (B) Secondary structure elements of the immunoglobulin C set (C=constant), which beta strands show the Greek key motif (RICHARDSON, 1981).*

F_{ab} fragments are composed of four Ig domains, two from the light chain, and two from part of the heavy chain, they are stabilized by intramolecular disulfide bonds (one in each Ig domain) and connected by one intermolecular disulfide bond. Because they are important molecules for many purposes e.g. immune therapy in inflammatory diseases and cancer (DEO *ET AL.*,

1997; WEINER, 1999), production processes with high yields are needed. A successful method of *in vitro* refolding of recombinantly expressed insoluble F_{ab} fragments in *E. coli* was described (BUCHNER AND RUDOLPH, 1991). The *in vitro* refolding of the four N-terminal Ig domains of the Axonin-1 from chicken showed, that this approach, which was used for f_{ab} fragments, can be modified successfully for immunoglobulin domains of neuronal cell adhesion molecules (FREIGANG *ET AL.*, 2000).

In general, crystallographic studies require large amounts of pure protein, which can be easily obtained by expression of insoluble inclusion bodies. Other more expensive methods are eucaryotic expression in yeasts, insect cells, or mammalian cells. Especially with mammalian cells, it can be difficult to obtain the required amounts, needed for subsequent protein crystallization experiments. It was shown for the transient axonal glycoprotein TAG-1 from chicken, and the E587-antigen$_{Ig1-4}$ from goldfish, that expression of both proteins, with subsequent in vitro refolding is successful (FREIGANG, 1999; FREIGANG *ET AL.*, 2000, THIS CHAPTER). In the case of chicken Axonin-1 and human TAG-1 both crystal structures hint to a homophilic interaction mechanism In the case of E587-antigen, crystals, which diffracted to 3.1 Å resolution, were obtained, but the structure was not solved, because phases could not be determined (FREIGANG, 1999).

B.1.8. Experimental goals

In order to examine the molecular basis of homophilic and heterophilic interactions several members of the L1 and the F11 family were investigated: Human TAG-1, gold fish E587-antigen, human L1, chicken NgCAM, chicken NrCAM, and chicken F11. The crystal structure of human TAG-1 was solved, and its implication on homophilic TAG-1 interaction is discussed in chapter C.

B.2. Material and Methods

B.2.1. Chemicals and laboratory equipment

Table B.1:
Suppliers of chemicals and equipment

2-mercaptoethanol	Fluka
Acetic Acid	Fluka
$AgNO_3$ (Silver nitrate)	Merck
Ampicillin	Sigma
Arginine free base	Fluka
Biotin	Fluka
Bradford	BioRad
BSA (Bovine serum albumin)	BioRad
Calcium chloride	Fluka
Chloramphenicol	Fluka
Coomassie R250	Sigma
Dextrose (Glucose)	Fluka
EDTA	Fluka
Ethanol	Fluka
Formaldehyde	Fluka
Glucose	Roth
Glutaraldehyde	Merck
Glutathione oxidized	Fluka
Gluthation reduced	Fluka
Glycerol	Fluka
Glycine	Fluka
HEPES	Fluka
Imidazole	Fluka
Kanamycin	Fluka
L-Cysteine	Fluka
lysine free base	Fluka
Methanol	Fluka

Table B.1:
Suppliers of chemicals and equipment

$MgCl_2$ x $6H_2O$ (Mangnesuim chloride)	Acros
Dried milk pouder	Migros (Switzerland)
Na acetate (Sodium acetate, anhydrous)	Fluka
Na_2CO_3	Merck
($Na_2S_2O_3$ x $5H_2O$) Sodium thiosulfate	Merck
NaCl (Sidium Chloride)	Riedel de Häen
NaOH (Sodium hydroxide)	Ridel de Häen
NZ-amines	Fluka
Peptone	Roth
Phosphoric Acid	Fluka
potassium phosphate	Sigma
(SDS) Sodium dodecyl sulfate	Roth
Sorbitol	Fluka
Tris (TRIZMA) etc	Sigma
Trypton	Roth
Urea	Fluka
Yeast Extract	Roth
(YNB) Yeast nitrogen base	Difco
Zeocin	Invitrogen

Table B.2
Buffers and solutions

TAE buffer (50x)	24,2% w/v Tris 5,7% v/v Acetic acid 10% v/v 0.5 mM EDTA pH 8.0 (NaOH)
TE buffer	1% v/v 1M Tris pH 8.0 0.2 % 0.5 mM EDTA pH 8.0 (NaOH)
Denaturing agent SB	5% w/v SDS 10% v/v 2-mercaptoethanol
Gel filtration buffer	20 mM Tris HCl pH 8.0 100 mM NaCl

Table B.2
Buffers and solutions

SDS-PAGE upper buffer (pH~6.7)	0.4 M Tris 0.4% w/v SDS
SDS-PAGE lower buffer (pH~8.8)	1.5 M Tris 0.4% w/v SDS
SDS-PAGE running buffer (10x)	3% w/v Tris 14.4% w/v Glycine 10% w/v SDS
Coomassie destaining solution	10% 2-Propanol
Coomassie staining solution	10% v/v Acetic acid X% w/v Coomassie R250
SDS-PAGE washing solution	10% Acetic acid 30% Ethanol
IEF cathode buffer (pH~10.2; 10x)	3.5% w/v Arginine (free base) 2.9% w/v Lysine (free base)
IEF anode buffer (pH~2.4; 50x)	4% v/v Phosphoric Acid (85%)
IEF sample buffer	20% v/v IEF Cathode Buffer 30% v/v Glycerol
Silver staining S1	30% v/v Ethanol 6.8% w/v Na-Acetate 0.2 % w/v $Na_2S_2O_3$ x $5H_2O$ 0.125 % Glutaraldehyde
Silver staining S2	0.2% w/v $AgNO_3$ 0.03% v/v Formaldehyde
Silver staining S3	2.5% w/v Na_2CO_3 0.02% v/v Formaldehyde
IMAC eqilibration buffer	8 M Urea 320 mM NaCl 100 mM Tris pH 8.5 20 mM 2-mercaptoethanol
IMAC washing buffer	8 M Urea 320 mM NaCl 100 mM Tris pH 8.5
IMAC elution buffer	8 M Urea 320 mM NaCl 250 mM Imidazole 100 mM Tris pH 8.5

Table B.2
Buffers and solutions

IMAC regeneration buffer	8 M Urea 2% v/v Acetic acid
IMAC elution buffer (native)	320 mM NaCl 250 mM Imidazole 100 mM Tris pH 8.5
Standard refolding buffer	0.8 M L-Arginine 0.32 M NaCl 100 mM Tris pH 8.5
Transformation buffer	100 mM $CaCl_2$ 70 mM $MgCl_2$ 40 mM Na-Acetate (anhydrous) pH 5.5 (Acetic Acid)

Table B.3
Media for E. coli and Pichia pastoris

LB	10 g Trypton 5 g Yeast Extract 7.5 g NaCl
Low Salt LB	10 g Trypton/Pepton 5 g Yeast Extract 5 g NaCl pH 7.5 (NaOH)
NZA	10 g NZ-amines 5 g Yeast extract 7.5 g NaCl
2YT	16 g Tryptone 10 g Yeast extract 5 g NaCl
YPD	10 g Yeast extract 20 g Peptone 20 g Dextrose (Glucose)
YPDS-Agar	10 g w/v Yeast extract 20 g w/v Peptone 20 g w/v Dextrose (Glucose) 10 g w/v Sorbitol

Table B.3
Media for E. coli and Pichia pastoris

BMMH/BMMG	100 mM potassium phosphate pH 6.0 1.34% v/v YNB 4 x 10-5% w/v Biotin 1% v/v Glycerol or 0.5% v/v Methanol

B.2.2. Strains, cells , vectors, and primers

B.2.2.1. *E. coli* strains

Different *E. coli* strains differ in expression and cloning properties. The strains which were used in this work can be classified according to their usage. Protease lacking expression strains were used to minimize the degradation of recombinantly expressed proteins (OmpT and lon were genetically inactivated). Strains for expression of constructs carrying the T7 promoter (pET and pTFT vectors) need the lysogenic variant of the λ-phage (DE3-strains), carrying the gene for the T7-RNA-polymerase, controlled by the lacUV5 promoter. To overcome problems due to the differences in codon usage of eucaryotic genes compared to *E. coli*, strains were used, which carry genes for the synthesis of rare tRNAs on an extra plasmid (pRIL, pRP, and pRILP strains). Expression experiments, using the pQE vectors do not need DE3-strains, because they carry a phage T5 promoter, controlled by two lac operator sequences to ensure efficient repression of the promoter, which allows the *E. coli* RNA polymerase to translate the DNA after IPTG induction, which efficiently displaces the repressor LacI from the operator (QIAGEN, 2003). In strains lacking the trxB gene alone or in combination with the gor gene, it is possible to express disulfide containing proteins within the cytoplasm, because both enzymes are involved in the maintenance of the native redox potential, which generally inhibits disulfide bond formation.

Table B.4
E. coli strains

Host strain	***Genotype***
E. coli BL21 (NOVAGEN, 2006)	*E. coli* B F^- *dcm ompT hsdS*(r_B^- m_B^-) *gal*
E. coli BL21 (DE3) (NOVAGEN, 2006)	*E. coli* B F^- *dcm ompT hsdS*(r_B^- m_B^-) *gal* λ(DE3)
E. coli BL21 (DE3) CodonPlus RP (STRATAGENE, 2005)	*E. coli* B F^- *dcm ompT hsdS*(r_B^- m_B^-) dcm^+ Tet^r *gal* λ(DE3) *end*A Hte [*argU proL* Cam^r]

Table B.4
E. coli strains

Host strain	***Genotype***
E. coli BL21 CodonPlus (DE3) pRIL (STRATAGENE, 2005)	*E. coli B F- dcm ompT hsdS*(r_B^- m_B^-) dcm^+ Tet^r *gal* λ(DE3) *end*A Hte [*argU IleY leuW* Cam^r]
E. coli BL21 CodonPlus (DE3) pRILP (STRATAGENE, 2005)	*E. coli B F- dcm ompT hsdS*(r_B^- m_B^-) dcm^+ Tet^r *gal* λ(DE3) *end*A Hte [*argU proL* Cam^r] [*argU IleY leuW* Strep/$Spec^r$]
E. coli Origami B (DE3) (NOVAGEN, 2006)	*E. coli* B F- *dcm ompT hsdS*(r_B^- m_B^-) Tet^r *gal* λ(DE3) *trxB gor*
E. coli BL21 (DE3) Rosetta gami 2 (NOVAGEN, 2006)	*E. coli B F- dcm ompT hsdS*(r_B^- m_B^-) Tet^r *gal* λ(DE3) *trxB gor* [*argU argN5 srgW argX glyT ileX leuW metT proL thrT thrU* Cam^r]
E. coli M15 (QIAGEN, 2003)	*E. coli* K12 Nal^S Str^S Rif^S Thi^- Lac^- Ara^+ Gal^+ Mtl^- F^- $RecA^+$ Uvr^+ Lon^+ [$laqI^q$ Kan^r]
E. coli SG13009 (QIAGEN, 2003)	*E. coli* K12 Nal^S Str^S Rif^S Thi^- Lac^- Ara^+ Gal^+ Mtl^- F^- $RecA^+$ Uvr^+ Lon^+ [$laqI^q$ Kan^r]
E. coli SB536 (BASS *ET AL.*, 1996)	*E. coli* K12 F^- WG1 ΔfhuA (tonΔ) ΔhhoAB (SacII)*shh*
E. coli W3110 (BACHMANN, 1972)	*E. coli* K12
E. coli TOP10 (INVITROGEN, 2004)	*E. coli* K12 F- *mcrA* Δ(*mrr-hsdRMS-mcrBC*) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G
E. coli XL10-Gold (STRATAGENE, 2006)	*E. coli* K12 Tet^R Δ(*mcrA183*) *end*A1 *supE44 recA1* Δ*(mcrCB-hsdSMR-mrr)173 gyrA96 relA1 lac Hte* [F^- *proAB* $lacI^qZ\Delta M15$ Tn*10* (Tet^R) Amy Cam^R]

B.2.2.2. P. pastoris

Recombinant expression of eucaryotic proteins in soluble form is often impossible in *E. coli* due to necessity of posttranslational modifications like disulfide bonding, glycosylation, phosphorylation, methylation etc. The yeast *P. pastoris* can be used to express recombinant proteins which need some of those posttranslational modifications. The system which was used in this work is based on the ability of *P. pastoris* to grow on methanol. Strains and vectors for *P. pastoris* expression were obtained commercially (INVITROGEN, 2005). Protocols for protein expression, Transformation, generating cryo-stocks etc. are used as described by the supplier. Only methods are described in detail, which differ from the supplier's manual.

Table B.5
Pichia pastoris strains

Strain	***Genotype***	***Phenotype***
X-33	wild-type	Mut^+
KM71H	*his4*	Mut^+, His^-
GS115	*arg4 aox1::ARG4*	Mut^S, Arg^+

(adapted from INVITROGEN, 2005)

B.2.2.3. Antibiotics

Antibiotics were used as follows: Tetracycline (12.5 mg/l), Kanamycin (50 mg/l), Chloramphenicol (34 mg/l), Ampicillin (200 mg/l), Zeocin (*E. coli*: 25 mg/l; *P. pastoris*: 100 mg/l), if other concentrations were not specified by the cells supplier.

B.2.3. Cloning

B.2.3.1. Colony PCR

E. coli or *P. pastoris* transformants were checked for correct insert length prior to sequencing, using the standard protocol of the Taq polymerase (NEW ENGLAND BIOLABS, 2005) with template DNA, prepared as follows:

P. pastoris

Transformants which were grown on agar-plates (approx. 10^5 cells) where transferred into 10 µl 0.5% w/v SDS. After vortexing (1 minute) and centrifugation (1 minute) 1-2 µl of the supernatant were used as template for PCR.

E. coli

Transformant colonies where transferred into 30 µl of ddH_2O and incubated at 95 °C for 2 minutes. After centrifugation (1 minute) 1-2 µl of the supernatant were used as template for PCR .

B.2.3.2. Mutagenesis

Experiments were done, using a commercially available kit (STRATAGENE, 2006). Using this method only one primer is needed, for each site. Components of the reaction buffer, which are not documented in detail in the suppliers manual, close the nicks between the elongated 5' ends and the neighboring primer's 3' end. A maximum of five sites, using one primer for each site, can be used for each mutagenesis experiment. Deviant to standard PCR, the amplified

products do not serve as templates in subsequent PCR cycles. The advantage is, that incorrectly synthesized PCR products are not amplified in subsequent PCR cycles. All clones, obtained after the following transformation are independently synthesized. For details of the whole mutagenesis procedure and the subsequent steps prior to cell transformation see the suppliers manual (STRATAGENE, 2006).

B.2.3.3. Restriction enzymes

Restricted digestion of PCR products and whole vectors were performed using the appropriate protocol, given by the suppliers manual (NEW ENGLAND BIOLABS, 2005).

B.2.3.4. Purification of DNA fragments

DNA fragments obtained by digestion with restriction enzymes from intact plasmid DNA and amplified PCR fragments were purified by electrophoresis using gels, containing 1.5% agarose, 0.05% ethidium bromide, and TAE buffer at constant voltage (80-120V) for 15-45 minutes, depending of the length of the used gel. DNA fragments were cut out manually, using UV light at a wavelength of 345 nm. The agarose pieces, containing the DNA fragments were purified using the Qiaquick Gel purification kit (QIAGEN, 2002). DNA fragments obtained by PCR, which were not subsequently digested with restriction enzymes, where purified with the PCR purification Kit from (QIAGEN, 2002).

B.2.3.5. DNA concentration

Determination of DNA concentration was done by measuring the extinction at 260nm (Eppendorf BioPhotometer), using an extinction coefficient for DNA of $\varepsilon = 50\ \mu g \cdot cm^{-1} \cdot ml^{-1.}$

B.2.3.6. Ligation

DNA fragments and vectors were ligated with a commercially available Kit. (ROCHE DIAGNOSTICS, 2004).

B.2.3.7. Chemically competent *E. coli* cells

Transformation of *E. coli* cells with plasmid DNA requires competent cells. For all transformations done with ligated or mutated plasmid DNA, commercially obtained chemical competent cells were used (*E. coli* TOP10, or *E. coli* XLGold see Table B.4, page 42). Only for transformation experiments done with intact plasmid DNA, chemical competent cells were used, prepared by the following procedure:

Competent cells

The desired *E. coli* expression strain, was grown 16 h in LB broth at 37 °C supplemented with the appropriate antibiotic. A 200 ml flask with 50 ml of medium and the appropriate antibiotic was inoculated with 0.5 ml of the overnight culture and grown to an OD_{600nm} of 0.3. Cells were centrifuged at 1000g in a 50 ml centrifugation vial for 1 min and resuspended in 10 ml transformation buffer, followed by a second centrifugation step. Prior to flash freezing in liquid nitrogen, cells were resuspended in 2 ml of transformation buffer supplemented with 7% DMSO and then transfered in 1.5 ml reaction vials in 100 µl aliquots. Competent cells were stored at -70 °C.

B.2.3.8. Transformation of *E. coli* cells

For each transformation reaction, a single aliquot of frozen chemically competent cells were thawed slowly on ice (0 °C) and 0.1-1 µl of plasmid DNA were added to the cells and mixed by tapping smoothly at the reaction vial. After 15 min the cells were transferred into a water bath (42 °C for 25 s), followed by incubation on ice for additional 2 minutes. After adding of prewarmed (42 °C) 0.5 ml 2YT medium, cells were shaken at 250 rpm in an incubator for 1 hour at 37 °C. Plating of the cells was done with glass beads and NZA plates adjuusted to 37 °C, supplemented with the appropriate antibiotic.

B.2.3.9. Sequencing

All constructs, which derived from a cloning procedure, which involved a PCR amplification step were sequenced using a sequencing service (GATC GmbH, Germany).

B.2.4. Protein analysis

B.2.4.1. Protein concentration

Spectroscopy

Using the primary sequence of a polypeptide chain, it is possible to calculate the extinction coefficient (Gill and von Hippel, 1989). The extinction at 280 nm is calculated by considering the number of tryptophane, tyrosine, and cysteine residues, with their specific extinction coefficients. In the strict sense, the calculation is valid only for denatured proteins in 6 M guanidine-HCl. A comparison with values obtained by a Bradford assay (Bradford, 1976) resulted in identical concentration values, thus showing a good correspondence between both methods for the proteins investigated in this work.

B.2.4.2. Electrophoresis

SDS-PAGE

To check the purity of protein samples at different stages of the purification process, it was found optimal to use 0.1-5 µg of the protein sample, depending on the staining technique. To increase the concentration of diluted samples (e.g. aliquots from refolding experiments or size exclusion chromatography) TCA precipitation was used:

TCA precipitation

The sample was mixed with the same volume of 20 % w/v TCA in 1.5 ml Eppendorf tubes, incubated on ice for 15 min, and centrifuged for 15 min at 4 °C at 10.000 g. The pellet was solubilized with reducing or nonreducing SDS sample buffer and incubated for 2 min at 95 °C prior to electrophoresis.

Standard sample preparation

Sufficiently concentrated samples were mixed with the same volume of 2 fold concentrated reducing or nonreducing sample buffer and treated like TCA precipitated samples prior to electrophoresis.

Refolding assay

Non reducing and reducing SDS-PAGE was used for analysis of the refolding experiments. Correctly refolded and oxidized samples run faster in electrophoresis, because the compactness of oxidized samples is enhanced, due to the intramolecular disulfide bonds. Prior to TCA precipitation of low concentrated samples, each sample was centrifuged for 10 min at 10.000 rpm to remove insoluble aggregates, and only the supernatant was used in subsequent sample preparation as described in the previous section.

Isoelectric focusing

To check the homogeneity of the samples with respect to their net surface charge, native low-salt (less than 50mM) or salt free samples were mixed with the same volume IEF sample buffer prior to electrophoresis. Running conditions: 1) One hour constant voltage 100 V; 2) One hour constant voltage 200 V; 3) 30 minutes constant voltage 500 V.

Coomassie staining

For sample detection of IEF and SDS-PAGE gels with protein amounts larger than 1 µg, the gels were incubated for 20 minutes in 20% ethanol and 10% acetic acid, followed by 10 minutes incubation in Coomassie solution at 85-100 °C. Destaining of the background was done with destaining buffer at room temperature until the protein bands were clearly visible.

Silver staining

For detection of protein ammounts smaller than 1 µg it silver staining was used. After electrophoresis, the gel was washed for 30 min in a mixture containing 20% ethanol and 10% acetic acid in ddH_2O. Then, the gel was incubated for 20 minutes in S1 buffer followed by five washing steps with ddH_2O for 2 min each. The silver reaction was then carried out with S2 buffer for 20 minutes, followed by washing two times with ddH_2O for 1 min each with vigorous shaking. Development of the gels was done by incubation with buffer S3 until band were clearly visible. The developing procedure was stopped with consecutive washing steps using ddH_2O.

B.2.5. Isolation of $L1_{Ig1-4}$, $NgCAM_{Ig1-4}$, $TAG\text{-}1_{Ig1-4}$, and $E587\text{-}antigen_{Ig1-4}$ from Inclusion bodies

B.2.5.1. Expression medium

To avoid basal protein expression prior to induction NZA medium was chosen for all expression experiments. Impurities within the LB-medium are able to induce like low concentration of IPTG.

B.2.5.2. Expression conditions

All expression experiments were done at 37 °C in a rotary shaker at 250 rpm. *E. coli* cells, transformed with one of the expression plasmids were grown for at least 16h with the appropriate antibiotic in NZA media until the OD_{600nm} reached a constant value. Frozen stocks were generated by adding DMSO to a final concentration of 7% prior to freezing 0.5 ml aliquots at -70 °C. Expression experiments were started with a preculture by inoculation of 200 ml NZA growth medium containing the appropriate antibiotic with one aliquot of the frozen cells in a 0.5 l baffled Erlenmeyer flask. At an OD_{600nm} of at least 1.0 the preculture was used to start the expression culture by inoculating 1l NZA aliquots containing the appropriate antibiotics in a 2l baffled Erlenmeyer flask with a starting OD_{600nm} of 0.01. After 2-4 hours protein expression was induced by 0.1-1 mM IPTG final concentration at an OD_{600nm} of 0.5-1. After 3.5 hours cells were harvested using a centrifuge at 10000g for 3 min. Cell pellets were slurried with deionized water to a final volume of 10 ml per liter expression culture and processed directly or stored at -70 °C for further use.

B.2.5.3. Isolation of inclusion bodies expressed in *E. coli*

For all expressed *E. coli* constructs a common procedure was used to obtain the respective purified denatured protein. Thawed or freshly centrifuged cell pellets were diluted with deionized water to 10 ml final volume, per liter expression culture. Cells were broken by to

passes through a constant cell disruption system at 2.5 kbar (Constant Systems Ltd., GB). Inclusion bodies were separated from cell debris and the soluble fraction by centrifuging the cells for 15 minutes at 10.000 g. Pelleted inclusion bodies were washed by three rounds of resuspending with deionized water and centrifugation for 15 minutes at 10.000 g. Inclusion body solubilization using denaturing solubilization buffer was carried out for at least 16 hours at room temperature using a magnetic stirrer. Prior to affinity chromatography, solubilized inclusion body protein was separated from non soluble components by centrifuging the sample at 75.000g for 1-2 hours.

B.2.5.4. Affinity chromatography of denatured inclusion body proteins

The supernatant obtained after inclusion body solubilization was loaded on an immobilized metal affinity chromatography (IMAC) column equilibrated with denaturing IMAC equilibrating buffer at a flow rate of 0.5-1 ml. After application of the sample, the column was washed with 3 column volumes (CV) of denaturing IMAC equilibration buffer, followed by 3 CV of denaturing IMAC washing buffer. Denaturing IMAC elution buffer was used for elution of the sample. All eluted fractions which contained protein were used for refolding experiments.

B.2.6. Refolding

B.2.6.1. General considerations

Recombinant protein expression in *E. coli* is often accompanied by solubility problems. Eukaryotic proteins, which fail to stay soluble in the prokaryotic host are often found as inclusion bodies. In this work inclusion body formation of the binding modules of different cell adhesion molecules were obtained, because for the stability of the binding modules, the postranslational formation of four disulfide bonds is necessary and because standard expression strains do not have an intracellular redox potential which allows disulfide bond formation in general. Oxidative refolding of denatured protein was successful for bovine pancreatic RNAse, which was shown almost 50 years ago (e.g. HABER AND ANFINSEN, 1961). However, the presence of cysteine residues, can lead to aggregation due to wrong intramolecular and intermolecular disulfide bonds as oxidation proceeds. To find a optimal in vitro refolding procedure, several parameters, which influence the refolding experiment have to be considered.

Minimization of aggregation

The denaturing buffer contained 8 M urea, which competes with the formation of native hydrogen bonds. To avoid aggregation during the refolding experiment upon urea dilution,

0.3 to 0.8 M L-arginine was added. L-arginine was successfully applied as inhibitor of aggregation in refolding experiments of F_{ab} fragments (BUCHNER AND RUDOLPH, 1991).

Redox system

As each of the binding module consists of four Ig domains, each stabilized by an internal disulfide bond, an appropriate redox system must be found for successful in vitro refolding. The binding modules tested here were purified under denaturing and reducing conditions to prevent wrong intramolecular and intermolecular disulfide bonds, which may result in aggregation of the target protein prior to refolding. As a starting point for the redox system, a combination of oxidized and reduced glutathione was used. Both components were shown to assisted correct disulfide bond formation of single chain F_{ab} fragments (BUCHNER AND RUDOLPH, 1991). Initial experiments with the proteins in this work showed, that variation of the redox system is essential to obtain at least a small fraction of fully oxidized protein. Therefore, systematic variation of the redox system was carried out.

Table B.6
Combinations of reducing and oxidizing agents tested for refolding

	DTT (0-3mM)	***Cysteamine (0.5-5mM)***	***GSH (0.2-2mM)***	***Cysteine (0.5-5mM)***
GSSG (0.5-5 mM)	●	●	●	●
Cystine (0.1 - 0.6 mM)	●	●	●	●
Cystamine (0.25-25 mM)	●	●	●	●

Overview of different redox systems used in refolding experiments. In parentheses are the concentration range, which was used in the respecting experiment. Each experiment covered 9 to 16 different redox conditions, which corresponded 3 to 4 different concentrations of reducing and oxidizing agent, respectively.

Refolding using dialysis

A slow removal of denaturing agent can be achieved by dialysis of the denatured sampled against a buffer, containing a lower concentration of the denaturing agent. A modification of a simple dialysis experiment is a so called controlled dialysis, where the concentration of denaturing agent in the dialysis buffer is reduced by controlled exchange of the dialysis buffer using a pump. This method was used successfully for refolding of reduced lysozyme (MAEDA *ET AL.*, 1995). The difference between a controlled and a standard dialysis experiment is the

rate of decrease of the denaturing agent, because it is proportional to the concentration difference between dialysis tube and dialysis buffer.

Refolding on a column

Affinity chromatography like IMAC (e.g. Ni-affinity chromatography) can be used to remove or dilute the denaturing agent faster than with dialysis. The protein is applied under denaturing conditions, and eluted under non denaturing conditions. A stepwise decrease in concentration of the denaturing agent or a linear decrease can be chosen. Compared to dialysis, the removal of the denaturing agent is much faster and the single molecules are separated from each other, due to their attachment to the column, which may reduce intermolecular interactions and thereby preventing aggregation during the refolding experiment.

Refolding using quick dilution

Fast removal of the denaturing agent is possible by dilution of the denatured sample with a buffer containing a lower concentration of the denaturing agent. An important parameter is a low final concentration of the target protein to avoid aggregation between folding intermediates due to hydrophobic interaction or due to intermolecular disulfide formation. This method is advantageous in cases where a screening of different conditions is necessary, e.g. examining the effect of different components of a redox system used in different ratios and combinations (see Table B.6, page 50). Small scale experiments can be done in parallel, whereas parallelization is almost impossible with affinity chromatography and difficult with dialysis.

Modification of the target

Systematic mutations of putatively surface-expressed protein residues (glutamate and lysine residues) has been shown to improve crystallization success and crystal lattice quality (DEREWENDA, 2004). In these examples exchanges for alanine often resulted in well diffracting crystals, but yielded less soluble protein. In this work an approach was developed were single hydrophobic residues were mutated to alanine, glutamic acid, or arginine. In a first step, the three dimensional structure of the target (in this work: $L1_{Ig1\text{-}4}$) was modeled using a web based server[1] (JAROSZEWSKI *ET AL.*, 2005). According to the obtained $L1_{Ig1\text{-}4}$ model, residues were selected, which have large hydrophobic side chains and were located on the surface of the model. Using the mutagenesis procedure described above, different $L1_{Ig1\text{-}4}$ mutations were generated (Table B.7, page 52). The procedure worked as follows: Day 1: PCR based mutagenesis reaction and subsequent generation of transformants in *E. coli* strain XL10-Gold. Day 2: inoculation of 5-10 precultures with colonies of the mutated transformants. Day 3: (a)

1 http://ffas.ljcrf.edu

Plasmid preparation from an aliquot (1 ml) of each preculture, and (b) expression experiments using the precultures to inoculate 20 ml volume of expression culture. A Coomassie stained gel of the expression cultures after 3 to 4 hours induction with IPTG (cultures were induced at an OD_{600nm} of 0.8 to 2.0) showed which mutants expressed $L1_{Ig1-4}$ with the predicted molecular mass (~42 kDa) and thus served as test, for positive clones. For conformation the plasmid DNA was sequenced. Expression cultures were frozen at -70 °C for further 3 to 4 hours after IPTG induction. Day 4: Isolation and purification of inclusion bodies from thawed day 3 cultures as described previously, with subsequent refolding experiments using the quick dilution method. Successfully mutated clones (verified by DNA sequencing) were used as template in subsequent mutagenesis trials to obtain $L1_{1-4}$ variants with several mutations in putative surface residues.

Table B.7
L1 primer used for solubility enhancement

Primer name	***Primer sequence***
L1_F95A	cccactctggctccGCcaccatcacgggcaac
L1_F104A_I111A	cagcaacGCtgctcagaggttccagggcGCctaccgctgc
L1_M172A_I176A	ggatctactggGCgaacagcaagGCcttgcacatcaagcagg
L1_L243A	agccgcgcctgGCcttccccaccaactcc
L1_Y290A	gccgaccgtgtcaccGCcccagaaccac
L1_L300A	cctgcagctgGCgaaagtgggcgaggaggatg
L1_L336A_L343A	ctgccccgtaGCggctgcacaagccccagagccatGCatatgggagg
L1_I371A	acctggagaatcaacgggGCccctgtggaggagctg
L1_F104R_I111R	cagcaacCGCgctcagaggttccagggcCGCtaccgctgc
L1_M172R_I176R	ctccggatctactggCGCaacagcaagCGCttgcacatcaagcagg
L1_L243R	ggaagccgcgcctgCGCttccccaccaactcc
L1_Y290R	gccgaccgtgtcaccCGCcagaaccacaacaagacc
L1_L300R	caagaccctgcagctgCGCaaagtgggcgaggaggatg
L1_L336R_L343R	ctgccccgtaCGCgctgcacaagccccagagccatCGCtatgggagg
L1_I371R	acctggagaatcaacgggCGCcctgtggaggagctg
L1_F95D	gccccactctggctccGATaccatcacgggcaacaac
L1_F104D_I111D	cagcaacGATgctcagaggttccagggcGATtaccgctgc
L1_M172D_I176D	ctccggatctactggGATaacagcaagGATttgcacatcaagcagg
L1_L243D	ggaagccgcgcctgGATttccccaccaactcc
L1_Y290D	gccgaccgtgtcaccGATcagaaccacaacaagacc
L1_L300D	caagaccctgcagctgGATaaagtgggcgaggaggatg

Table B.7
L1 primer used for solubility enhancement

Primer name	***Primer sequence***
L1_L336D_L343D	ctgccccgtaGATgctgcacaagccccagagccatGATtatgggagg
L1_I371D	acctggagaatcaacgggGATcctgtggaggagctg

Only those putatively surface exposed residues were considered in the solubility enhancement experiments, which do not overlap with identified natural mutations, as they are supposed to be implicated in altered L1 homophilic and heterophilic binding properties (see Figure B.3).

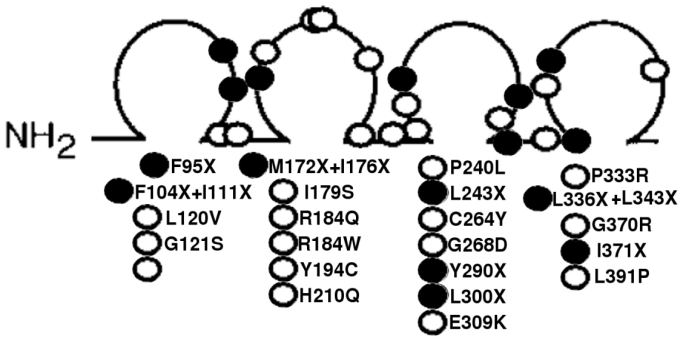


Figure B.3: Schematic picture of the four N-terminal Ig domains of human L1. *Each Ig domain is symbolized by a loop. The relative position of natural mutations involved in L1 disease are indicated as unfilled circles, target positions of residues, which are chosen for point mutations in order to enhance the refolding process are indicated as filled circles. The figure is modified from the original publication (KENWRICK AND DOHERTY, 1998).*

B.2.6.2. Refolding of $L1_{Ig1-4}$ and $NgCAM_{Ig1-4}$ - Experimental procedures

Refolding using dialysis

With $L1_{Ig1-4}$ and $NgCAM_{Ig1-4}$ refolding experiments based on dialysis were carried out. $L1_{Ig1-4}$ or $NgCAM_{Ig1-4}$ samples denatured and prepared with 6M guanidine HCl instead of 8M urea were given into a 50 to 100 µl dialysis tube and dialyzed against standard refolding buffer, containing 0.5 M guanidine HCl instead of 0.8 M L-arginine. Variations were done by changing oxidizing and reducing agent concentrations.

Refolding on a column

With $L1_{Ig1-4}$, $NgCAM_{Ig1-4}$, and TAG-1_{Ig1-4} IMAC based refolding experiments were carried out. As in the first step of affinity chromatography of denatured inclusion body proteins (chapter B.2.5.4.) 1 mg or less protein was bound on a 1 ml column. Instead of elution with denaturing buffer, the column was washed with non denaturing IMAC washing buffer (3 CV) and subsequently eluted with non denaturing IMAC elution buffer and analyzed by non reducing SDS-PAGE.

Refolding of $L1_{Ig1-4}$ and $NgCAM_{Ig1-4}$ using quick dilution

The standard small scale refolding procedure was carried out in 1.5 ml Eppendorf test tubes, using 40 µg of $L1_{Ig1-4}$ or $NgCAM_{Ig1-4}$ dissolved in a volume of 10-40 µl of denaturing IMAC elution buffer supplemented with the reducing agent. The reduced protein mixture was attached to the inner face of the Eppendorf test tube's cap and mixed with 1 ml refolding

buffer by vortexing the test tube for 30 s. Non standard experiments with larger volumes were carried out as similar as possible to the standard small scale refolding procedure. Only the dilution and mixing procedure was therefor deviated from the standard small scale refolding procedure: vortexing was replaced by shaking the aliquots by hand. All refolding aliquots were stored at 4 °C.

Refolding of modified $L1_{Ig1-4}$

All refolding experiments of modified $L1_{Ig1-4}$ were carried out in 1.5 ml Eppendorf test tubes (according to the standard small scale refolding procedure), using 40 µg of mutated L1 dissolved in a volume of 10-40 µl of denaturing IMAC elution buffer supplemented with 4 µM of DL-cysteine mixed by vortexing for 30 s with 1 ml standard refolding buffer and incubation at 4 °C. Aliquots of the refolding samples were checked by non reducing SDS-PAGE as described in chapter B.2.4.2.

B.2.7. *Pichia pastoris* expression experiments

Yeast expression is based on the ability of *Pichia pastoris* to grow on methanol as sole carbon source. *Pichia pastoris* possess two genes AOX1 and AOX2 which encode for two different alcohol oxidases which are involved in the first step of methanol degradation. Using the pPICαA vector the gene of interest is integrated in the host's genome by recombination after the AOX1 promoter which is responsible for methanol induced expression. For protein expression, a frozen stock of *P. pastoris* recombinants were used to inoculate 100ml of growth medium in a 0.5 l baffled Erlenmeyer flask at 30 °C and 400 rpm until an OD_{600nm} of at ~5 was reached. The preculture was then centrifuged at 3000 rpm and the pellet was resuspended in 500ml BMM (BMMH for strain GS115) with a starting OD_{600nm} of 1.0 using 2l baffled Erlenmeyer flasks. Every 12 hours 0.5% of methanol was added to the expression cultures. After 3-5 days the supernatant was separated from the cells, filtered sterile, and stored at 4°C for further use.

B.2.8. HEK cell expression

Expression tests of the constructs in HEK 293 (with L1 pEAK8 vectors) cells were done at the University of Zürich in the group of Prof. Peter Sonderegger by Dr. Kerstin Leuthäusser. But, for both constructs, with native L1 and Calsyntenin-1 (VOGT *ET AL.*, 2001) signal sequence, it was not possible to detect any expression in the supernatant of the cells.

B.2.9. Deglycosylation of *Pichia pastoris* expression products

Secreted expression of recombinant proteins in *P. pastoris* is advantageous compared to *E. coli*, because it is easier to obtain recombinant protein product with proper disulfide bonding. But a postranslational modification like glycosylation can be a problem for subsequent crystallization experiments, due to the flexibility of the sugar chains. *P. pastoris* adds in average 8-14 mannose units on the respective asparagine side chain (N-glycosylation), which is much less compared to the yeast *Saccharomyces cerevisiae* (50-150 mannose units). To remove or truncate the the oligosaccharide chains after expression, two different enzymes were used: N-Glycosidase F (Roche and New England Biolabs) and Endoglycosidase H (New England Biolabs). The reaction products of N-Glycosidase F are ammonia, the complete oligosaccharide, and aspartic acid in the peptide chain, which is the only modification of the protein chain. In contrast, Endoglycosidase H does not completely remove the oligosaccharide, but cuts between the two innermost N-acetylglucosamine units, attached to the polypeptide chain, thus one N-acetylglucosamine unit remains attached to the asparagine residue. All experiments were done with a 5 to 10 fold molar excess of deglycosylation enzyme compared to predicted glycosylation sites (GUPTA *ET AL.*, 2004). Glycosylation site prediction was done using a web based interface[#].

Deglycosylation under denaturing and native conditions condition

Prior to deglycosylation a small volume of the protein sample was dialyzed against the corresponding reaction buffer as described by the supplier. The deglycosylation reaction was performed as recommended by the supplier, with a 1 to 5 fold higher enzyme concentration as given in the original protocol.

http://www.cbs.dtu.dk/services/NetNGlyc/

B.3. Results and Discussion

As described in the introduction, the structure of the binding module of Axonin-1 was solved by FREIGANG *ET AL.*, 2000. TAG-1 the human ortholog of Axonin-1 shows a similar spatial orientation of the first N-terminal Ig domains, but an unexpected specific crystal contact hints to an alternative binding mode of the human ortholog compared with chicken Axonin-1. The first four Ig domains of TAG-1 were expressed as insoluble inclusion bodies, and refolded by a quick dilution method derived from a protocol described for single F_{ab} fragments (BUCHNER AND RUDOLPH, 1991). The TAG-1 expression, purification and its three dimensional structure is described in chapter C. The binding modules of all investigated cell adhesion molecules are responsible for the majority of homophilic and heterophilic interactions. The strategy, to start with several target molecules instead of focusing on only one molecule, was chosen to overcome problems during expression and refolding. Furthermore it provided a basis to study the molecular basis of heterophilic interactions between the recognition modules from different molecules e.g. human L1 and TAG-1, chicken TAG-1 and NrCAM, or F11 and NgCAM.

B.3.1. Constructed vectors

B.3.1.1. Already constructed *E. coli* vectors

The structure of chicken Axonin-1 and the methods used by FREIGANG *ET AL.*, 2000 served as a starting point. For five different cell adhesion molecules, an expression construct of the binding module was constructed earlier, using the vectors pTFT and pQE60 (MÖRTL, 2000), namely the following constructs: human TAG-1 (TAG-1-pTFT and TAG-1-pQE60), human L1 (L1-pTFT and L1-pQE60), chicken NgCAM (NgCAM-pTFT and NgCAM-pQE60), chicken NrCAM (NrCAM-pTFT and NrCAM-pQE60), and chicken F11 (F11-pTFT and F11-pQE60). Details like primer design and restriction sites used for cloning are given in the appendix. To determine the optimal length of all constructs, a multiple alignment of all target molecules and Axonin-1 was made. The final constructs were made, such that only few additional residues were added at the N and C-terminal parts of the binding module. Hydrophobic amino acids where also avoided (see Figure B.4).

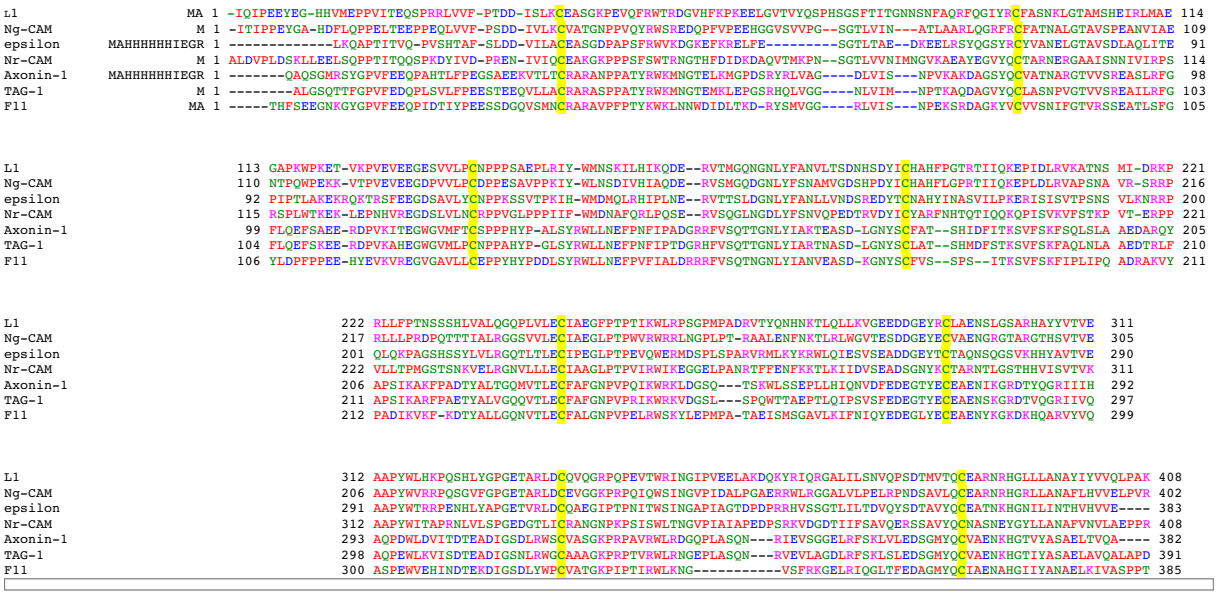


Figure B.4: Alignment of the seven target cell adhesion molecules. *Each of the four blocks is equivalent to one Ig-domain as identified by the structures of chicken Axonin-1 and human TAG-1.*

B.3.1.2. pET15b expression vectors

TAG-1

To increase the low level of TAG-1 expression, achieved with TAG-1-pTFT (2 mg/l as inclusion bodies) a pET15b (NOVAGEN, 2006) derived expression construct was cloned. A PCR amplified fragment of the TAG-1 cDNA was ligated into the multi cloning site of the pET15b vector, such that the length of the expressed polypeptide chain was identical as in the case of TAG-1-pQE60 or TAG-1-pTFT. C-terminal hexa-histidine tag (HisTag) and stop codon were added during the PCR amplification.

F11 and NgCAM

To test a different vector for expression of F11 and NgCAM, two pET15b derived expression constructs were cloned, using an approach identical to TAG-1-pET15b construction. In addition, for both targets expression constructs, with N-terminal HisTag were made. Primers and vector sequences of F11-NHis-pET-15b, F11-CHis-pET-15b, NrCAM-NHis-pET-15b, and NrCAM-CHis-pET-15b are given in the appendix.

B.3.1.3. Modified L1

The following modified L1-pQE-60 constructs were successfully cloned sequenced and expressed in *E. coli* strain SG13009.

Figure B.5

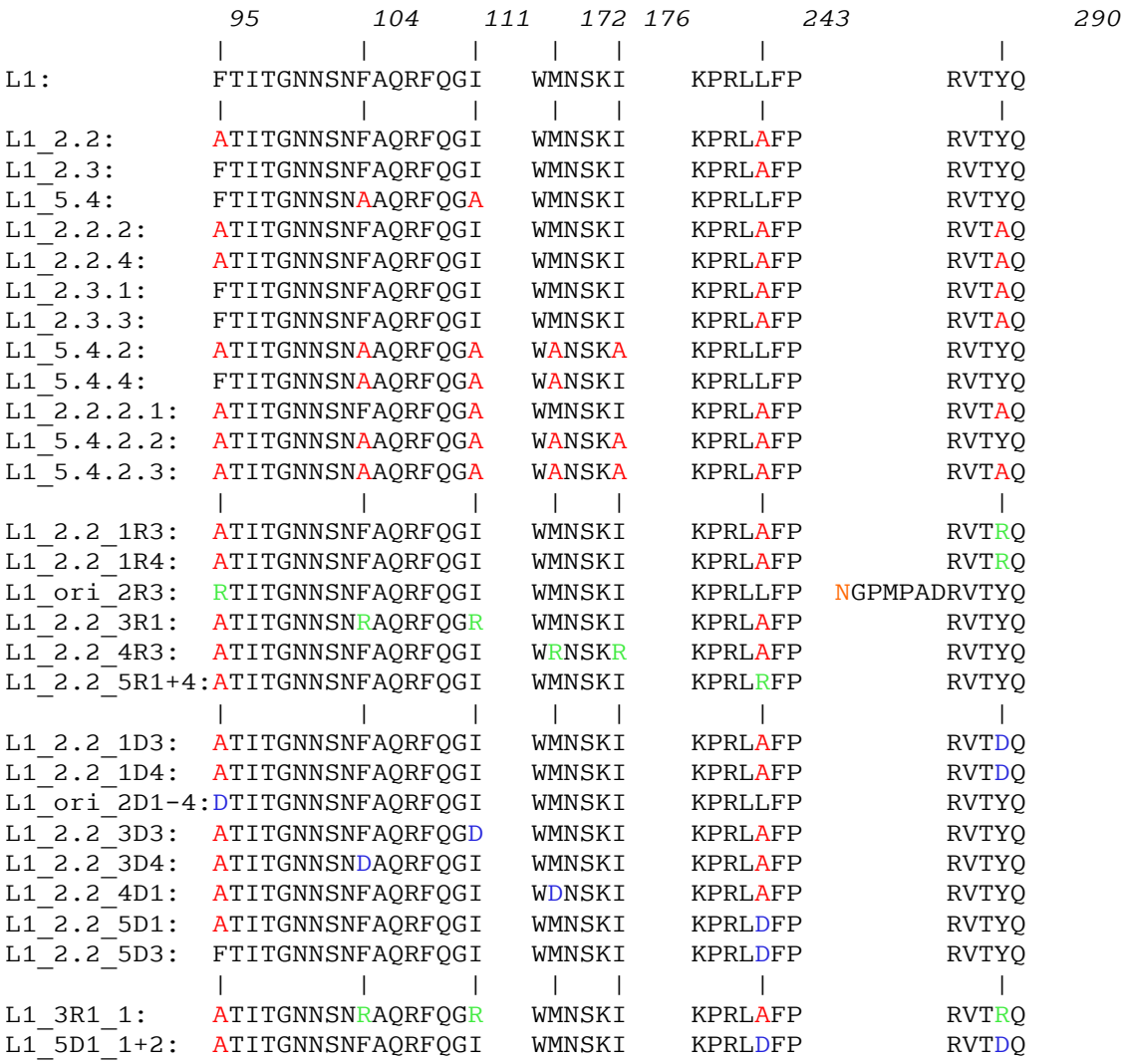


Figure B.5: L1 mutants of the L1-pQE-60 construct obtained by site directed mutagenesis. ***Red****: alanine mutants to decrease the size of the corresponding hydrophobic amino acid.* ***Blue****: aspartic acid mutants to increase solubility by a negatively charged residue.* ***Green****: arginine mutants to increase solubility by a positively charged residue. All shown constructs were verified by sequencing and were expressed as inclusion bodies and used for refolding experiments.*

B.3.1.4. *P. pastoris* expression vectors

General considerations

Construction of expression vectors for secreted expression in the yeast *P. pastoris* were made according to the supplier's manual (INVITROGEN, 2005). Length and position of the

HisTag were not modified, compared to the *E. coli* expression constructs. The encoded region of all constructs is elongated only by the α-signal sequence of the pPICZαA vector.

L1-pPICZaA and TAG-1-pPICZaA

Cloning procedure, DNA sequence and position of the insert according to the α-signal cleavage site are given in the appendix.

E587-antigen-pPICZaA

For E587-antigen a special *P. pastoris* construct was made. It was constructed with an N-terminal hexa HisTag and a factor Xa cleavage site, to cut off the HisTag prior to crystallization. The construct was shortened N- and C-terminally, compared to a previously refolded and crystallized variant (FREIGANG, 1999). Both ends of the expressed polypeptide chain were shortened according to the multiple alignment (see Figure B.6.), such that the the first and the last residue correspond to to the first and the last residue, which were ordered in the structure of Axonin-1. Sequencing of the expression vectors showed, that the published sequence (GIORDANO *ET AL.*, 1997) differs from that used for expression. The position of the missing 8 cysteine residue was identified, which is responsible for correct disulfide bonding within the fourth Ig domain (see Figure B.6).

Figure B.6
DNA sequence of E587-antigen

```
Published       ctaaaacaagctccaacgatcacagttcagcccgtgtctcacactgctttctctctggatgatgtgattctggcc 75
Sequenced       ctaaaacaagctccaacgatcacagttcagcccgtgtctcacactgctttctctctggatgatgtgattctggcc 75
                 L  K  Q  A  P  T  I  T  V  Q  P  V  S  H  T  A  F  S  L  D  D  V  I  L  A
                 L  K  Q  A  P  T  I  T  V  Q  P  V  S  H  T  A  F  S  L  D  D  V  I  L  A

Published       tgtgaagccagtggagacccggcacccagcttccgctgggtgaaagatggcaaagagtttaagagggaactgttg 150
Sequenced       tgtgaagccagtggagacccggcacccagcttccgctgggtgaaagatggcaaagagtttaagagggaactgttt 150
                 C  E  A  S  G  D  P  A  P  S  F  R  W  V  K  D  G  K  E  F  K  R  E  L  L
                 C  E  A  S  G  D  P  A  P  S  F  R  W  V  K  D  G  K  E  F  K  R  E  L  F

Published       agttcgggaactctgacagcagaagataaagaggagctcc-atcctatccagggatcgtaccgctgctacgtg--- 221
Sequenced       gagtctgggaccctgacagcagaagataaagaggagctccgatcctatc-agggctcgtaccgctgctacgtggcc 225
                 S  S  G  T  L  T  A  E  D  K  E  E  L   H  P  I  Q  G  S  Y  R  C  Y  V
                 E  S  G  T  L  T  A  E  D  K  E  E  L  R  S  Y   Q  G  S  Y  R  C  Y  V  A

Published       -atgagcttgggaaccgccgtctcagacctggcgcagctgatcacagagcccatcccaaccttggccaaagagaag 297
Sequenced       aatgagct-gggaaccgccgtctcagacctggcgcagctgatcacagagcccatcccaaccttggccaaagagaag 300
                   M  S  L  G  T  A  V  S  D  L  A  Q  L  I  T  E  P  I  P  T  L  A  K  E  K
                 N  E  L   G  T  A  V  S  D  L  A  Q  L  I  T  E  P  I  P  T  L  A  K  E  K

Published       aggcagaaaacaagatcgtttgaggagggagacagcgcagtcctctactgtaaccctcccaagagctctgtcact 372
Sequenced       aggcagaaaacaagatcgtttgaggagggagacagcgcagtcctctactgtaaccctcccaagagctctgtcact 375
                 R  Q  K  T  R  S  F  E  E  G  D  S  A  V  L  Y  C  N  P  P  K  S  S  V  T
                 R  Q  K  T  R  S  F  E  E  G  D  S  A  V  L  Y  C  N  P  P  K  S  S  V  T

Published       cccaaaatacactggatggacatgcactggcgccacatccctctgaatgagcgggtgaccactagtcttgatgga 457
Sequenced       cccaaaatacactggatggacatgcagctgcgccacatccctctgaatgagcgggtgaccactagtcttgatgga 450
                 P  K  I  H  W  M  D  M  H  W  R  H  I  P  L  N  E  R  V  T  T  S  L  D  G
                 P  K  I  H  W  M  D  M  Q  L  R  H  I  P  L  N  E  R  V  T  T  S  L  D  G

Published       aatctgtactttgccaacttactagtcaatgacagcagagaagactacacctgcaatgcccacatcatcaacgcc 522
```

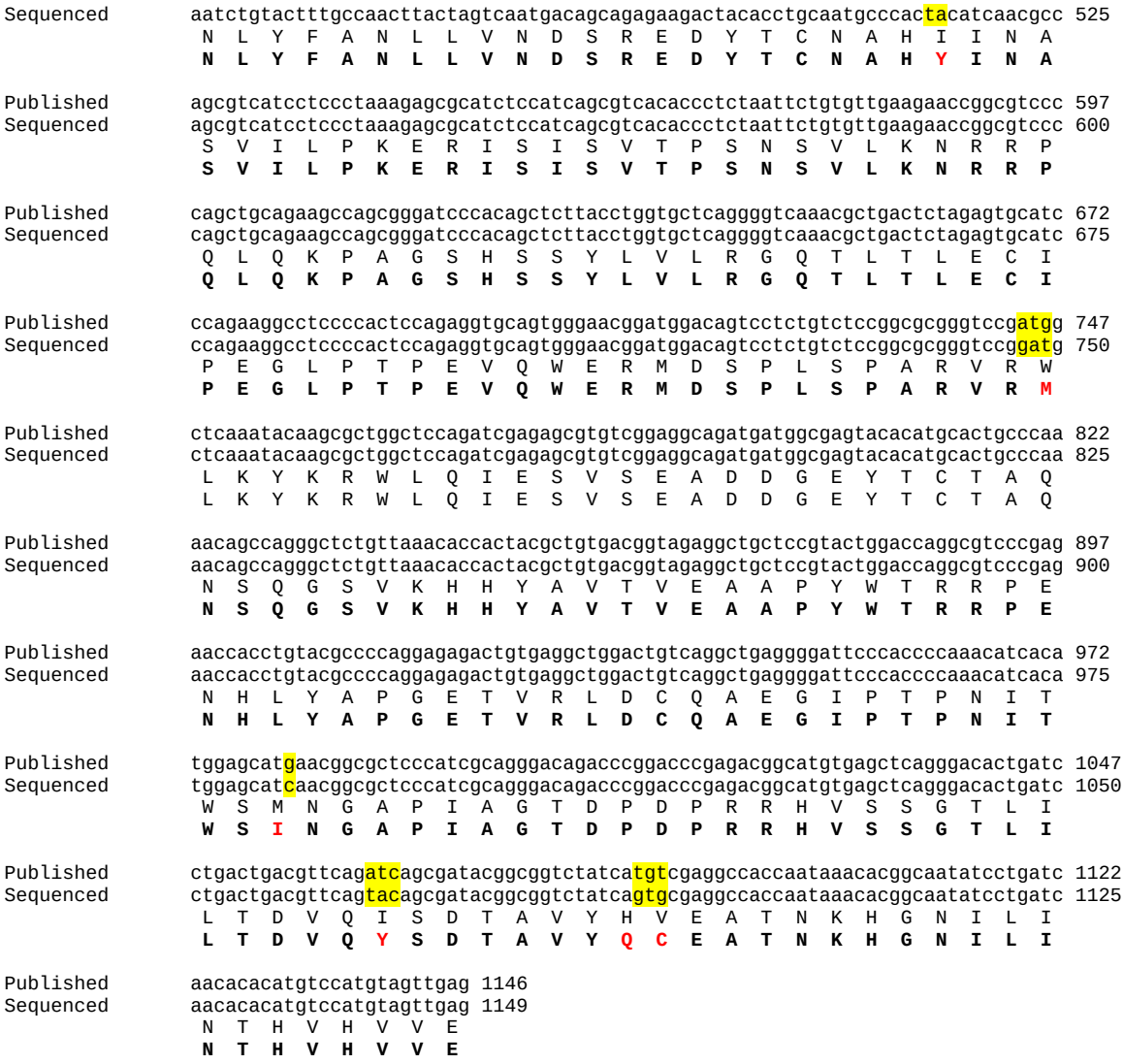


Figure B.6: DNA and amino acid sequence alignment of the E587-antigen*. The Published sequence is from the database as published by (*GIORDANO *ET AL.*, 1997*). The Sequenced sequence is as sequenced from two independently constructed E587-antigen-pPICZαA vectors. Differences at the DNA level are highlighted in yellow, and deduced differences at the amino acid level are highlighted in red.*

B.3.1.5. Vectors for HEK293 cell expression

Two constructs for human L1 where made for expression in human embryonic kidney cells (Cell line HEK293). One of the constructs had a native L1 signal sequence as secretion signal, an other had the signal sequence of Calsyntenin-1 (VOGT *ET AL.*, 2001), but both constructs did not lead to successful $L1_{Ig1\text{-}4}$ secretion small scale expression cultures (both vector sequences are given in the appendix).

B.3.2. *E. coli* expression experiments of $L1_{Ig1-4}$, $TAG\text{-}1_{Ig1-4}$, $NgCAM_{Ig1-4}$, $NrCAM_{Ig1-4}$ and $F11_{Ig1-4}$, Axonin-1_{Ig1-4}, and E587-antigen$_{Ig1-4}$

Initial *E. coli* expression experiments were done earlier, but only L1 and TAG-1 had been expressed reproducibly in amounts sufficient for further purification and refolding experiments (MÖRTL, 2000). A systematic approach, using different *E. coli* strains and varying types of growth medium and induction time was chosen for bacterial expression experiments.

B.3.2.1. Vectors and *E. coli* strains

Table B.8
Vectors and strains, successfully used for inclusion body expression

Target protein	Vector	E. coli strain
$TAG\text{-}1_{Ig1-4}$	[1]**TAG-1-pET-15b** [2]TAG-1-pTFT	BL21 (DE3) pRILP
$L1_{Ig1-4}$	[2]L1-pQE60	SG13009
$NgCAM_{Ig1-4}$	[1]**NgCAM-pTFT**	BL21 (DE3) pRIL
E587-antigen$_{Ig1-4}$	[3]E587-antigen-pTFT	BL21 (DE3)
$NrCAM_{Ig1-4}$	[1]**NrCAM-Chis-pET-15b** [1]**NrCAM-Nhis-pET-15b**	BL21 (DE3) pRILP
$F11_{Ig1-4}$	[1]**F11-Chis-pET-15b** [1]**F11-Nhis-pET-15b**	BL21 (DE3) pRILP
Axonin-1 (as control)	[3]Axonin-1-pTFT	BL21 (DE3)

[1]These vectors were successfully are expressed in this work.
*[2]These vectors were already successfully expressed earlier (*MÖRTL, 2000*).*
*[3]These Vectors were already successfully expressed earlier (*FREIGANG, 1999*).*

B.3.2.2. Isolation of $L1_{Ig1-4}$

$L1_{Ig1-4}$ isolated from inclusion bodies

Successful expression experiments were carried out earlier for L1 using L1pQE60 vector with *E. coli* strains SB536, W3110, SG13009 and M15 (induction with IPTG) or with L1-pTFT and *E. coli* strains BL21, BL21 pRP, SB536, and W3110 (induction with the phage CE6), for details see (MÖRTL, 2000). Refolding experiments were carried out, using purified inclusion bodies obtained from SG13009 with L1-pQE60. 20 mg denatured L1 purified from inclusion bodies per liter *E. coli* expression culture could be obtained reproducibly.

$L1_{Ig1-4}$ isolated from the cytoplasm

No soluble $L1_{Ig1-4}$ was obtained from the cytoplasm of *E. coli* strains BL21 (DE3), BL21 (DE3) pRIL, and BL21 (DE3) pRILP.

B.3.2.3. Isolation of TAG-1_{Ig1-4}

TAG-1_{Ig1-4} isolated from inclusion bodies

Successful expression of TAG-1 was shown earlier for TAG-1-pTFT with *E. coli* strains BL21 (DE3) and BL21 (DE3) pRIL (MÖRTL, 2000). Expression of TAG-1 was possible in *E. coli* strain BL21 (DE3) pRILP too. The additionally constructed TAG-1-pET-15b expression vector, was used successfully to express TAG-1 as inclusion bodies. Both constructs yielded 1-2 mg purified denatured TAG-1_{Ig1-4} per liter *E. coli* expression culture.

TAG-1_{Ig1-4} isolated from the cytoplasm

No soluble TAG-1_{Ig1-4} was isolated from the cytoplasm of *E. coli* strains BL21 (DE3), BL21 (DE3) pRIL, and BL21 (DE3) pRILP.

B.3.2.4. Isolation of $NgCAM_{Ig1-4}$

Earlier attempts to produce the NgCAM inclusion bodies were successful but not reproducible (MÖRTL, 2000). Using the *E. coli* strain BL21 DE3 (pRIL) it was possible to obtain reproducibly 5 mg purified denatured $NgCAM_{Ig1-4}$ per liter expression culture. The difference to earlier expression experiments was the use of NZA medium instead of LB medium.

B.3.2.5. Isolation of $F11_{Ig1-4}$

Earlier attempts to produce F11 inclusion bodies were not successful (MÖRTL, 2000). Using the *E. coli* strain BL21 DE3 (pRILP) and the F11-pET-15b vectors it was possible to obtain reproducibly 5 mg purified denatured $F11_{Ig1-4}$ per liter expression for both constructs, with C-terminal or N-terminal HisTag..

B.3.2.6. Isolation of $NrCAM_{Ig1-4}$

Earlier attempts to produce $NrCAM_{Ig1-4}$ inclusion bodies were not successful (MÖRTL, 2000). Using the *E. coli* strain BL21 DE3 (pRILP) and the NgCAM-pET-15b vectors, it was possible to obtain reproducibly 5 mg purified denatured $NgCAM_{Ig1-4}$ per liter expression for both constructs, with C-terminal or N-terminal HisTag.

B.3.2.7. Isolation of E587-antigen$_{Ig1-4}$

Using the already successfully used E587-antigen-pTFT vector (FREIGANG, 1999) for expression in *E. coli* strain BL21 (DE3), it was possible with NZA medium to obtain 5 mg purified E587-antigen$_{Ig1-4}$ per liter expression culture

B.3.2.8. Isolation of Axonin-1$_{Ig1-4}$

As a control for refolding experiments, Axonin-1 was expressed successfully as reported (FREIGANG, 1999). Deviant from the previously conducted experiments, NZA expression medium was used instead of LB medium. 25 mg denatured Axonin-1$_{Ig1-4}$ was purified per liter expression culture.

B.3.2.9. All target molecules could be obtained as inclusion bodies

Although the first trials with pQE-60 derived expression constructs did not result in the expression of all five new targets TAG-1, L1, F11, NrCAM, and NgCAM, but only in expression of L1 (MÖRTL, 2000), further trials with the expression vector pTFT resulted in successful overexpression of TAG-1 (MÖRTL, 2000) and NgCAM (this work). Finally, pET derived expression constructs resulted in successful overexpression of TAG-1 and C- and N-terminally His-tagged F11 and NrCAM. To change non-expressing constructs into expressing constructs DNA sequence changes to overcome codon bias problems has been found useful (GUSTAFSSON *ET AL.*, 2004), but in this work it was sufficient to change the expression vector. Some background information about promoters of different expression vector systems are explained in a review (BANEYX, 1999).

B.3.3. *In vitro* folding

Solubilization of proteins expressed as inclusion bodies is possible with buffers containing high concentrations of denaturing agents like urea or guanidine-HCl. In vitro folding of such denatured proteins is a trial an error process, because only few predictions can be derived from the primary structure, concerning effective refolding conditions in terms of refolding rate and correct tertiary structure. Refolding of targets containing predominantly beta sheets is probably more difficult than alpha helical targets, because the residues, which are involved in hydrogen bonding can be far apart in the primary sequence, whereas the hydrogen bonds in alpha helices infer residues which are only four residues apart. Disulfide bonds, which are required for stability of the correctly refolded target have to be formed by oxidation of two cysteine residues. In this work, the largest effect in the in vitro refolding rate was observed by variations of the redox system, which is required to facilitate correct formation of disulfide bonds and simultaneously inhibits the formation of non-native ones (wrong intra- and intermolecular disulfide bonds, which may result in aggregates).

B.3.3.1. Refolding of $L1_{Ig1-4}$ and $NgCAM_{Ig1-4}$

Refolding using dialysis

The experiments were done with $L1_{Ig1-4}$ and $NgCAM_{Ig1-4}$. L1 and NgCAM protein samples were taken as eluted from the IMAC column (6 M guanidine HCl, 100 mM Tris HCl, 500 mM NaCl, 250 imidazole, pH 8.5, Protein concentration 5 mg/ml), and diluted to a Protein concentration of 150 mM with dialysis buffer (3 M guanidine HCl, 50 mM Tris HCl, 1 mM ETDA, 200 mM NaCl, 400mM L-arginine, 50 mM DTT, pH 9.0). Dialysis was carried out in four steps in modified dialysis buffer without DTT but with 0.4 mM oxidized glutathione and descending guanidine HCl concentration (2 M, 1 M, 0.5 M, 0 M) at 4 °C. Dialysis buffer was exchanged every 12 hours. Only very faint bands of oxidized L1 and NgCAM were detected with Coomassie staining as shown in Figure B.7.

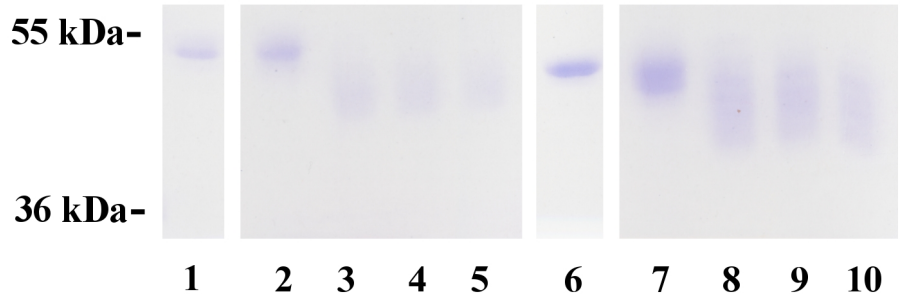


Figure B.7: Stepwise dialysis of the four N-terminal Ig domains of human L1 and chicken NgCAM concentration. *From left to right: Lane 1: L1 reduced, in dialysis buffer containing 2 M guanidine HCl. Lane 2-5: L1, non reduced, decreasing guanidine HCl (2 M, 1 M, 0.5 M, 0 M). Lane 6 to 10 correspond to the same experiment using NgCAM.*

Refolding on a column

The target molecule was bound to the IMAC column (1ml bed volume) in denatured conformation, and eluted after removal of the denaturing agent (8 M urea or 6 M guanidine-HCl) with IMAC elution buffer (native). Assaying the eluted samples with non-reducing and reducing SDS-PAGE, showed that only 10-20% of the initial sample amount were eluted, but without disulfide formation, as the reducing and non reducing sample showed no difference in apparent molecular mass. Non eluted sample was eluted only by IMAC regeneration buffer under denaturing conditions. Additional experiments, where the denaturing buffer was removed gradually, with a 30 ml gradient mixer prior to elution, did not

result in correctly refolded and oxidized L1 or NgCAM. These results suggest that L1 and NgCAM disulfide formation is a process which cannot occur on the short time scale of an IMAC based separation experiment as used here.

Refolding using quick dilution

Most experiments were done using the quick dilution method (see page 51). Several combinations and variations of the redox system tried in this work, showed that oxidation of the cysteines is the critical step, because many conditions were identified, where $L1_{Ig1-4}$ or $NgCAM_{Ig1-4}$ were oxidized non uniformly: Five distinct refolding products were identified using the refolding assay in an apparent molecular weight range between 55 and 40 kDa (see Figure B.8).

Each of the five distinct bands may correspond to different $L1_{Ig1-4}$ or $NgCAM_{Ig1-4}$ products, varying by the amount of oxidized cysteines. As the largest band has an identical apparent molecular mass as the reduced control sample (not shown), it may correspond to non oxidized $L1_{Ig1-4}$ or $NgCAM_{Ig1-4}$. Descending from the band with the second largest apparent molecular weight to the smallest one, one disulfide bond may be present additionally. Only the smallest band, may therefore represent the correctly refolded and oxidized target, which contains four

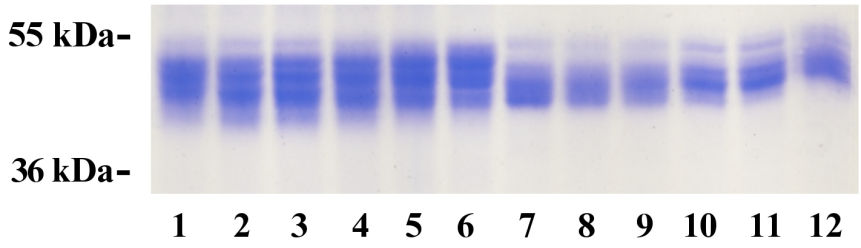


Figure B.8: Screening of different redox conditions of L1 and NgCAM refolding. *The fife distinct band can be identified best in lanes 2-4. Lane 1-6: L1 refolding in 100 mM Bis-Tris-Propane, pH 9,9, 100 mM NaCl, 0.5 M L-arginine, 1 mM EDTA, 3 mM DTT and increasing GSSG concentration: 0.5 mM, 1 mM 2 mM 4 mM, 8 mM 16 mM. Lane 7-12: Refolding of NgCAM, using identical conditions as for L1.*

disulfide bonds, one in each Ig domain. The results from the refolding assay are not sufficient, to rule out that the molecules of one band are identical. That mean it is not sure that e.g. the second highest band is formed by molecules which each have the disulfide bond formed within the same Ig domain, or if randomly one of the four domains possess one disulfide bond. The formation of distinct bands may however suggest, that the disulfide bonds, which are formed are native, because wrongly formed disulfide pairs of cysteines originating from

different Ig-domains would perhaps disturb the regular pattern. This possibly occurred in the refolding experiments using dialysis (Figure B.7).

Isolation of refolded $L1_{Ig1-4}$

Chromatographic purification of the product, forming the band with the smallest apparent molecular weight was not successful for $L1_{Ig1-4}$ or $NgCAM_{Ig1-4}$.

Purification with the following chromatographic separation media were tested, according to the recommended standard protocol of the respective supplier: Ion exchange, using (a) Poros 20HQ (Applied Biosystems), (b) Source 30 Q, (c) Resource Q , (d) Mono Q, (e) chromatofocusing using Mono P, (f) affinity chromatography using activated thiol sepharose 4B (b-f: GE Healthcare). But all experiments showed, that the difference of the postulated L1 species (varying with respect to their disulfide bond content) have only weak influence on the properties responsible for chromatographic separation.

B.3.3.2. Refolding of TAG-1_{Ig1-4} and E587-$antigen_{Ig1-4}$.

The refolding procedure for both expressed constructs was similar and the successful conditions are described in chapter C. TAG-1 and E587-$antigen_{Ig1-4}$ refolding is finished after complete oxidation of all cysteines within two weeks. A shift in apparent molecular weight from approximately 50 kDa to 40 kDa was considered as a criterion for refolding success. In contrast to L1 and NgCAM, only two distinct bands could be observed, corresponding to the fully oxidized and the reduced sample. There was no fraction with an intermediate apparent molecular weight as observed in the refolding experiments with L1 and NgCAM.

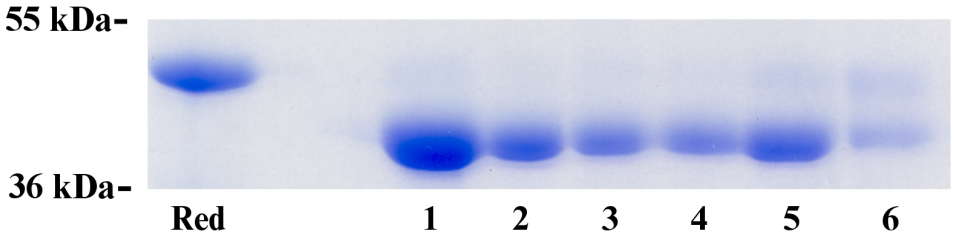


Figure B.9: Non reduced gel of refolded human TAG-1. *The six different refolding experiments (1-6) are from protein samples, which are expressed and IMAC purified independently. The gel shows that the refolding procedure is very reproducible. Red.: Reduced TAG-1 for comparison.*

The refolding rate (two weeks for complete oxidation) was comparable to that, which was reported for Axonin-1 (six weeks, FREIGANG, 1999). However, in control experiments with Axonin-1 under the same refolding conditions as used here (L-cysteine as the only component of the redox system), Axonin-1 showed an increased refolding rate: complete oxidation within two days for small scale experiments (1 ml) and two weeks as observed for TAG-1 and E587-antigen in large scale experiments (100 ml).

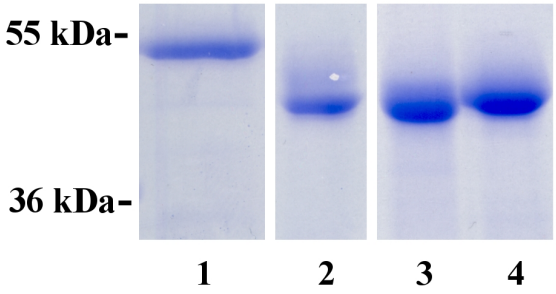


Figure B.10: Refolding of E587-antigen. *Lane 1 and 2: Reduced and non-reduced sample of a refolding experiment. Lane 3 and 4: Reduced refolding sample after buffer exchange in factor Xa reaction buffer. Lane 3 is after 16 h factor Xa digestion, lane 4 without digestion. Western blot with His-Tag specific antibodies showed no signal in factor Xa treated samples (not shown).*

B.3.3.3. Refolding of $F11_{Ig1-4}$ and $NrCAM_{Ig1-4}$

For both expressed constructs no optimal refolding condition was found, which resulted in a homogeneous and uniform band with a smaller apparent molecular mass. In contrast to TAG-1 and E587-antigen L-cysteine as the only component of the redox shuffle system was not sufficient for correct $F11_{Ig1-4}$ and $NrCAM_{Ig1-4}$ refolding. The most successful

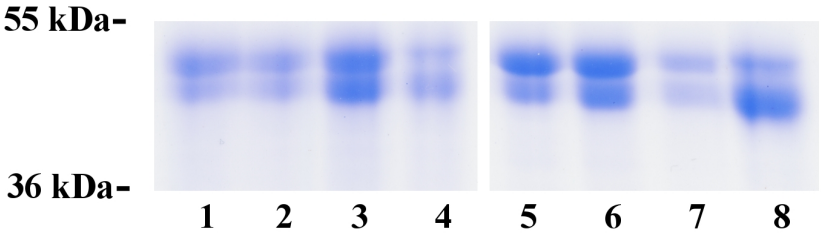


Figure B.11: F11 refolding after 5 weeks*. Left half: F11 with N-terminal His-Tag; right side: C-terminal His-Tag. Refolding was done with a redox system of 4mM oxidized glutathione and descending DTT concentration (3 mM, 1 mM, 0.5 mM, 0 mM, from left to right.)*

concentrations of the components of the redox system for $F11_{Ig1-4}$, which however resulted in incomplete oxidation, were 4 mM oxidized glutathione in combination with 0-3 mM DTT as reducing agent (Figure B.11). The F11 refolding species can be compared with TAG-1 or E587-antigen refolding products, because only one oxidized species could be detected, however it was not possible to find conditions, which resulted in complete oxidation of F11.

In the case of NrCAM, incomplete refolding was observed using L-cysteine as only component of the redox system in the refolding experiment. Only a small fraction was oxidized even after twelve weeks (see Figure B.12). The refolding products of NrCAM resembled the pattern observed in L1 and NgCAM refolding but with a less discrete distribution. In contrast to the refolding product of TAG-1 and E587-antigen, more than two species (a completely reduced and a completely oxidized one) could be detected.

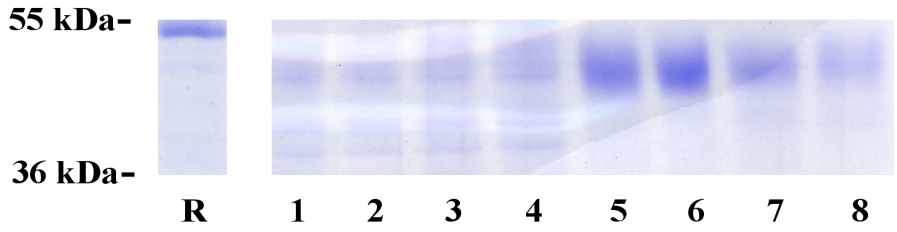


Figure B.12: Refolding of NrCAM. *Lane 1: NrCAM reduced. Lane 2-5: Refolding of C-terminally with increasing L-cysteine His-Tagged NrCAM. Lane 6-9: N-terminally His-Tagged NrCAM. From left to right(PAGE 58)*

B.3.3.4. The different refolding success may resemble target similarities

According to the refolding success the target molecules used here can be divided into two different groups. $TAG\text{-}1_{Ig1-4}$, $E587\text{-}antigen_{Ig1-4}$, $Axonin\text{-}1_{Ig1-4}$, and $F11_{Ig1-4}$ as representatives of group one which can be successfully refolded, and $L1_{Ig1-4}$, $NgCAM_{Ig1-4}$, and $NrCAM_{Ig1-4}$ as members of group2, which resisted refolding in all experiments carries out. $F11_{Ig1-4}$ can be classified as member of the successful group, because it shows the characteristic two species in the refolding experiments, which correspond to the reduced and the completely oxidized species of the target. In contrast, all members of group two $L1_{Ig1-4}$, $NgCAM_{Ig1-4}$, and $NrCAM_{Ig1-4}$, show more than two species in the refolding assay. The sequence similarity between these three targets is high (L1-NgCAM [58%], L1-NrCAM [41%], NgCAM-NrCAM [40%]), and may be the reason for similar behavior in refolding experiments. By electron microscopy using negative staining, it was possible to show, that $L1_{Ig1-4}$ is present in two different conformations: in a compact U-shaped conformation similar to that observed for TAG-1, Axonin-1, and hemolin and alternatively in an extended conformation (SCHÜRMANN *ET*

AL., 2001). Additionally, it was observed that L1 is able to form intermolecular disulfide bonds, and it was speculated that this disulfide bond formation is mediated by the first two Ig domains (HASPEL *ET AL.*, 2001). Although the latter observation was made with constructs comprising only the three N-terminal Ig domains, the general ability to form intermolecular disulfide bonds may cause the missing success in *in vitro* refolding experiments of the second target group (L1, NgCAM, and NrCAM). Studies with refolding of different single Ig domains showed that stability and refolding rate of this structural class of protein domains is highly variable (CLARKE *ET AL.*, 1999). This variability of refolding rate between different domains, together with the need for disulfide bridge formation within each single domain of the targets used in this work, may result in refolding intermediates, which are trapped in a certain state, where the molecules are refolded partly. If each Ig-domain needs a different redox shuffle system for successful disulfide bridge formation, it may be impossible to find a condition where cysteine oxidation can occur in all four Ig domains.

B.3.3.5. Purification of TAG-1_{Ig1-4}, and E587-$antigen_{Ig1-4}$ after refolding

The purification of TAG-1_{Ig1-4} is described in chapter C. A modified version of the TAG-1_{Ig1-4} protocol was used for E587-$antigen_{Ig1-4}$: The fully oxidized protein from the refolding aliquots, were concentrated, by ultrafiltration using an Amicon ultrafiltration cell (Millipore Inc.) with 10.000 kDa molecular weight cutoff (YM10) and Viviascience 6ml concentrator (MWCO: 10 kDa, Vivascience GmbH) 200 fold. Concentrated E587-$antigen_{Ig1.4}$ was subjected to size exclusion chromatography (Superdex 200, GE Healthcare) using a buffer containing 20 mM Tris-HCl pH 8.0, 150 mM NaCl, and 10% w/v glycerol and a flow rate of 1-2 ml/min. Without glycerol as stabilizing agent, the fraction of monomeric E587-antigen was almost completely converted to multimeric aggregates, which were eluted within the exclusion volume of the column.

B.3.3.6. Cleavage of the N-terminal HisTag of E587-$antigen_{Ig1-4}$ after refolding

After a concentration procedure as described in the previous section, E587-antigen was subjected to buffer exchange chromatography (1ml NAP5 column, GE Healthcare) incubated with factor Xa reaction buffer. The proteolytic cleavage of the N-terminal HisTag was then carried out for 16 hours at 25 °C. For each ml refolding aliquot (~40µg E587-$antigen_{Ig1-4}$) 1µg factor Xa was used. Cleaved and non-cleaved E587-$antigen_{Ig1-4}$ were separated by IMAC batch purification slurry as provided by the supplier. The proteolytic cleavage was successful, but cleaved and non-cleaved E587-$antigen_{Ig1-4}$ precipitated easily in the factor Xa reaction buffer at concentrations smaller than 2 mg/ml.

B.3.4. *Pichia pastoris* expression experiments

An approach to express the target molecules in soluble form was the use of *P. pastoris*. The expression constructs were chosen, that expressed target molecules were secreted into the growth medium. These expression experiments were conducted to obtain correctly folded target proteins with all disulfide bonds established. All genetic constructs were tested by PCR for their Mut phenotype prior to expression experiments, as recommended and described by the supplier (INVITROGEN, 2005), because differences in the Mut phenotype can lead to different expression levels and different expression conditions. The MutS phenotype looses the integrity and functionality of the AOX1 gene during integration of the target insert into the the hosts genome. In an analytical PCR experiment, with one primer specific for the AOX1 promoter and a second primer specific for the 3' region of the AOX1 gene the Mut^+ phenotype can be characterized by two amplified DNA products, the AOX1 gene and the inserted target

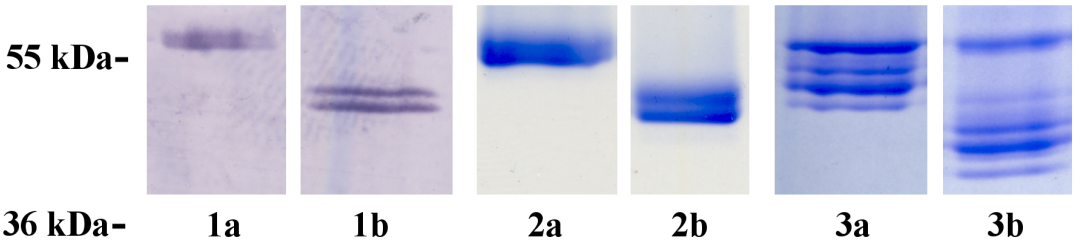


Figure B.13: Native and deglycosylated TAG-1 as expressed by Pichia pastoris*. (1) Western blot prior (a) and after (b) deglycosylation under denaturing and reducing conditions. (2) Coomassie stained gel with TAG-1 samples prior (a) and after (b) deglycosylation under reduced and denaturing conditions. (3) TAG-1 samples prior and after deglycosylation under native and non reducing conditions. Sample (a) prepared with reducing (a) and non reducing (b) sample buffer.*

gene, whereas in Mut^S mutants no DNA product corresponding to the AOX gene can be detected. Additionaly the Mut^S phenotype can be verified in an growth rate experiment, because the missing AOX1 gene is responsible for decrease on growth rate, as only the gene product of the AOX2 gene can be used for ethanol degradation. Due to their genetic background, KM71H strains are Mut^S even without insert integration . Details are described in the suppliers manual (INVITROGEN, 2005).

B.3.4.1. *P. pastoris* expression of TAG-1 and L1

Successful overexpression conditions were found for TAG-1 and L1 in Pichia strain X-33, using the standard expression procedure as given in the supplier's manual. Both clones

secreted the target molecule into the growth medium (BMMY). TAG-1 expression was induced by addition of 0.5 % methanol every 12 hours, whereas for L1 expression 1.5 % methanol was used. Expression of TAG-1 was confirmed by Western blot using an HisTag specific primary antibody. Compared to *E. coli P. pastoris* expressed TAG-1 showed two species, which differed in apparent molecular mass; both were larger than *E. coli* expressed TAG-1. After deglycosylation under denaturing conditions, the apparent molecular mass of TAG-1 is smaller and in the range of its counterpart from *E. coli*, but two species remained. TAG-1 detection was possible in SDS-PAGE with Coomassie staining as well as with Western blot. As the two different sized species remained after deglycosylation, it is very likely, that both species differ, because of two different cleavage sites of the signal sequence (see Figure B.13). This problem is discussed in the suppliers manual and was confirmed by personal communication with the manufactures technical service. It can only be solved by generating new genetic constructs, varying within the N-terminal target sequence.

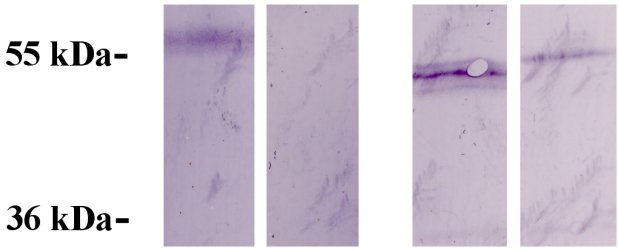


Figure B.14: Western blot identification of TAG-1 an L1 in the P. pastoris supernatant with HisTag specific antibodies*. Left side: Detected TAG-1 and non detected L1 prior to deglycosylation with PNGase-F. Right side: Human TAG-1 as well as L1 can be detected after the reducing and denaturing deglycosylation procedure. Samples were reduced, prior to electrophoresis.*

L1 expression in *P. pastoris* did not result in a expression product in the supernatant, which could be detected by Western blot prior to deglycosylation. However after deglycosylation under denaturing conditions, L1 was identified with the same apparent molecular mass as *E. coli* expressed L1 inclusion bodies. Although the L1 is supposed to be completely denatured after SDS-PAGE and the blotting procedure, the attached sugar moieties are responsible for this result; by an unknown mechanism they interfered or blocked antibody binding.

Although the L1 fraction in the *P. pastoris* supernatant was soluble, in contrast to TAG-1, the L1 fraction seemed to be in a non native conformation. This conclusion was made, because after ion exchange chromatography (with Poros 20 HQ), and non reducing deglycosylation the reduced samples ran predominantly as a single band on the gel, whereas, non reduced

samples showed much larger aggregates, which originated from intermolecular disulfide bonds (Figure B.15).

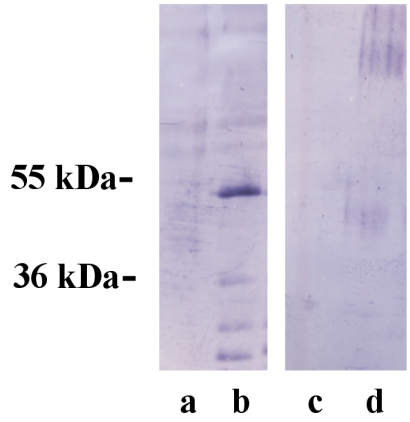


Figure B.15: L1 oligomerization by intermolecular disulfide bonds were detected by Western blot. *Comparison of reduced (a+b) and non reduced (c+d) L1. Before deglycosylation (a+c) L1 is not detectable. Only after PNGase-F treatment under non reduced, native conditions (b+d) the HisTag specific primary antibody has bound to blotted L1.*

B.3.4.2. Difficulties with *P. pastoris* expression

TAG-1$_{Ig1-4}$

TAG-1 expression in P. *pastoris* was successful with respect to amount and disulfide content as detected by SDS-PAGE and Western blot. The different species which were identified differed in N-terminal signal sequence processing and were detected as two separate bands. Although ion exchange chromatography seemed to be the separation procedure of choice, it was not possible to separate both bands from each other. As refolding experiments conducted in parallel, were successful, no further investigation on *P. pastoris* expressed TAG-1 were carried out.

L1$_{Ig1-4}$

Although other cell adhesion molecules were expressed successfully in *P. pastoris* (Ig 1 of NCAM: KISELYOV *ET AL.*, 1997; IG 1-2-3 OF NCAM: SOROKA *ET AL.*, 2003) L1 expression resulted in expression species crosslinked by intermolecular disulfide bonds. Additionally the amount of L1 expression was very low compared to TAG-1 expression, as L1 detection after deglycosylation was possible only with Western blot. Western blot analysis also showed large

intermolecular cross linked L1 expression products. As discussed previously in the refolding section, disulfide bond mediated dimerization was observed in vitro (HASPEL *ET AL.*, 2001), as well as an alternative elongated L1 conformation of the four N-terminal Ig domains (SCHÜRMANN *ET AL.*, 2001), in addition to the U-shaped variant as seen in hemolin, Axonin-1 and TAG-1 structure. If those results hint to a native feature of L1, related to modulation or

Figure B.16: L1 Purification from P. pastoris with anion exchange chromatography*. The silver stained gel of the elution profile shows that L1 is not pure after this chromatography step. The Western blot after reducing and denaturing deglycosylation of L1 shows that L1 was bound to the column. This chromatography step using a POROS 20 HQ column with 2 ml total bed volume was introduced to concentrate L1 from 50 ml of supernatant.*

regulation of growing axons in the developing nervous system, *P. pastoris* expression as well as refolding experiments might have failed because the structural variability probably was too large for both systems.

C. The crystal structure of the ligand binding module of human TAG-1 suggests a new mode of homophilic interaction

This work is in preparation for submission:

MÖRTL, M., DIEDERICHS, K., SONDEREGGER, P., WELTE, W. (2006) The crystal structure of the ligand binding module of human TAG-1 suggests a new mode of homophilic interaction.

C.1. Abstract

Human TAG-1 is a neuronal cell adhesion molecule that is crucial for the development of the nervous system during embryogenesis. It consists of six immunoglobulin-like and four fibronectin-III-like domains anchored to the membrane by glycosylphosphadidylinositol. Here we present the crystal structure of the four N-terminal immunoglobulin like domains of TAG-1 (TAG-1_{Ig1-4}), known to be important in heterophilic and homophilic receptor interaction. The contacts of neighboring molecules within the crystal were investigated, and a comparison with the structure of the chicken ortholog resulted in two alternative models for the molecular mechanism of homophilic TAG-1 interaction. Both models of TAG-1 homophilic interaction, proposed in this paper are based on dimer interaction and not on formation of a molecular zipper as proposed for the chicken ortholog.

C.2. Introduction

Neuronal cell adhesion molecules serve as ligand and receptor for molecular guidance and play a crucial role during the development of the nervous system. These molecules are involved in axon growth, path finding, and axon guidance along their predetermined pathways (Tessier-Lavigne and Goodman, 1996). Neuronal cell adhesion molecules are expressed on the cellular surface and particularly on the growth cone of axons, but they can be found also as components of the extracellular matrix or in soluble form in the environment of growing axons (Dodd and Jessell, 1988b; Jessell, 1988). Axonal cell adhesion molecules (AxCAMs) have been grouped into different families on the basis of the domains in their extracellular part. Several important members of this family like TAG-1, L1, and NgCAM are composed of repeated domains of immunoglobulin superfamily (IgSF) and fibronectin-III-like (FnIII) domains (Chothia and Jones, 1997). They act as sensor molecules on the developing axon, recognizing and transducing growth signals from the environment by mediating adhesive cell-cell and cell-matrix interactions (Brümmendorf and Rathjen, 1996; Hynes and Lander, 1992; Rathjen and Jessell, 1991). The human transient axonal glycoprotein TAG-1 belongs tom the F11-family of AxCAMs which all have in common six N-terminal Ig and four FnIII domains attached by a glycosylphosphatidylinositol anchor to the cell membrane (Hasler *et al.*, 1993; Tsiotra *et al.*, 1993). Other members of the F11-family are F11/F3/contactin, BIG-1, BIG-2, NB-2, and NB-3 (Brümmendorf *et al.*, 1989; Gennarini *et al.*, 1989; Kamei *et al.*, 1998; Ogawa *et al.*, 1996; Ranscht, 1988; Yoshihara *et al.*, 1995).

Orthologs of human TAG-1 (also known as TAX1) are found in mouse and rat and in chicken (axonin-1) (Dodd *et al.*, 1988; Stoeckli *et al.*, 1989; Yamamoto *et al.*, 1986). An interesting feature of TAG-1 is the maintenance of their binding features across the species

border. The characteristics of heterophilic interaction of human TAG-1 with human L1 (described later) is similar to the heterophilic interaction of the mouse and chicken orthologs with the corresponding L1 orthologs (NgCAM). This functional similarity between different species was exploited in several studies, where heterophilic binding was studied between binding partners originating from different species, e.g. interaction of human TAG-1 with rat L1 or chicken TAG-1 with human L1 (De Angelis *et al.*, 1999; Milev *et al.*, 1996; Pavlou *et al.*, 2002; Tsiotra *et al.*, 1993). This indicates that these molecules have not lost or changed their function in the development of the nervous system during evolution between humans, rodents and birds allowing the direct comparison of results, obtained from experiments with orthologous molecules.

Other important members of the IgSF family are the transmembrane proteins L1, NgCAM and NrCAM, which belong to the L1 family (Burgoon *et al.*, 1991; Moos *et al.*, 1988). Their common characteristic is a fifth FnIII-domain, a transmembrane helix, and a small intracellular domain (Grumet *et al.*, 1991; Grumet and Edelman, 1984; Lemmon *et al.*, 1989). Many members of the IgSF family of AxCAMs are known to act both as a receptor and as a ligand. The interaction of the binding partners can occur in the same membrane (*cis*) and across the intracellular space from two molecules located on different cells (*trans*).

TAG-1 was initially described and purified as an axonally secreted protein from the dorsal root ganglia neurons (Stoeckli *et al.*, 1989), and is transiently expressed during the development of central nervous system and peripheral nervous system as a glycosylphosphatidylinositol-anchored and as a secreted form (Dodd *et al.*, 1988; Karagogeos *et al.*, 1991) which is the predominant form found in chicken (Stoeckli *et al.*, 1991). During the development of the human nervous system, TAG-1 expression is very specific and can be utilized as a marker of axonogenesis (Karagogeos *et al.*, 1997). Recent studies have shown, that TAG-1 is expressed by corticofugal axons, where it serves as a substrate for migrating cortical interneurons, and that its expression is regulated by thyroid hormone (Alvarez-Dolado *et al.*, 2001; Denaxa *et al.*, 2001). TAG-1 has also been implicated in axon-glia interactions, and is probably involved in the tumorigenesis of malignant gliomas (Rickman *et al.*, 2001; Suter *et al.*, 1995; Traka *et al.*, 2002).

Apart from the ability to interact heterophilically with other AxCAMs and with extracellular matrix compounds, TAG-1 and its orthologs are able to interact homophilically. Chicken TAG-1 expressed on myeloma cells as well as human TAG-1 expressed on S2 cells have the ability to induce cell-cell aggregation by a *trans* interaction of TAG-1 (Kunz *et al.*, 2002; Rader *et al.*, 1993; Tsiotra *et al.*, 1996). The crystal structure of chicken TAG-1 revealed important information to narrow down the region on the four N-terminal Ig-domains, responsible for homophilic TAG-1 interaction (Freigang *et al.*, 2000). The analysis of the crystal packing together with site directed mutagenesis experiments resulted in a zipper like

model for the homophilic interaction of TAG-1 in aggregating myeloma cells (FREIGANG *ET AL.*, 2000).

The crystal structure of human TAG-1 which is presented here hints to an alternative mode of homophilic *trans* interaction, which is equally in accordance with the mutagenesis results of (FREIGANG *ET AL.*, 2000).

It is known that TAG-1 can act as binding partner in homophilic and heterophilic interactions. Heterophilic interactions of TAG-1 with L1, NgCAM, and NrCAM were investigated extensively. *In vitro* studies showed, that neurite outgrowth on an NgCAM substrate is blocked by antibodies against NgCAM or L1 but not by antibodies against TAG-1 (BUCHSTALLER *ET AL.*, 1996; FELSENFELD *ET AL.*, 1994; KUHN *ET AL.*, 1991). These findings, together with additional *in ovo* studies (STOECKLI AND LANDMESSER, 1995), showed that TAG-1 does not interact with NgCAM in *trans,* which is in agreement that Schneider line 2 (S2) cells expressing either TAG-1 or L1 do not form mixed cell aggregates (BUCHSTALLER *ET AL.*, 1996; MALHOTRA *ET AL.*, 1998). In neurite outgrowth experiments heterophilic interaction of TAG-1 with NgCAM was only observed in *cis* (BUCHSTALLER *ET AL.*, 1996; STOECKLI *ET AL.*, 1996). However, an NrCAM substrate is known to induce neurite outgrowth by a heterophilic *trans* interaction with TAG-1 (LUSTIG *ET AL.*, 1999). *In ovo* studies proved the importance of the heterophilic *trans* interaction of floor plate cell NrCAM with TAG-1 of neurites of dorsal root ganglia (LUSTIG *ET AL.*, 1999; STOECKLI AND LANDMESSER, 1995). TAG-1 interaction with NrCAM is also involved in the formation of contacts between glial cells and neurites. Other heterophilic interaction partners of TAG-1 are F11/F3, β-integrin, Neurocan, Phosphacan/RPTP-ζ/β, and NCAM (BUTTIGLIONE *ET AL.*, 1998; FELSENFELD *ET AL.*, 1994; MILEV *ET AL.*, 1996). Responsible for NgCAM and NrCAM binding are the four N-terminal immunoglobulin like (Ig) domains of TAG-1 (FITZLI *ET AL.*, 2000; KUNZ *ET AL.*, 1998), which form a structural entity called the ligand binding module, that can maintain its structural and functional integrity only in the presence of all four N-terminal domains (this work and (FREIGANG *ET AL.*, 2000; RADER *ET AL.*, 1996)).

C.3. Experimental procedures

C.3.1. Expression and purification of TAG-1_{Ig1-4}

DNA coding for the four N-terminal Ig-domains of TAG-1_{Ig1-4} (residues 28-418) was cloned into the vector pET15b (Novagen). The construct contained one extra N-terminal amino acid (M) and eight extra C-terminal amino acids containing a hexa-histidine tag (RSHHHHHH). Expression was carried out in *Escherichia coli* strain BL21 (DE3) pRILP (Novagen) using a medium containing 10 g N-Z-amine A (Sigma), 5 g Yeast extract (Fluka), and 7.5 g NaCl

(Sigma) per liter at 37 °C. Protein expression was induced at an optical density of 0.9 at 600nm with 1 mM isopropyl-β-D-thiogalactopyranoside (Fermentas). Three hours after induction cells were harvested by centrifugation at 10000 g for three minutes, and cell pellets were resuspended in distilled water and broken using a continuous cell disruption system (Constant Systems Ltd., GB), at 2.5 kbar. Inclusion bodies were washed by three cycles of centrifugation and resuspension with distilled water and solubilized by overnight incubation with 8 M Urea (Sigma), 100 mM Tris-HCl (Fluka) pH 8.5, 320 mM NaCl, 50 mM 2-mercaptoethanol (Fluka) (solubilization buffer) at room temperature. The solute was centrifuged for 2 hours at 75000 g and the supernatant was filtered through a 0.22 µm filter and purified using a 4 mL HisTrap HP column (GE Healthcare) using the following protocol at a flow rate of 0.5 ml/cm²: 1) Column equilibration with 3 column volumes (CV) of solubilization buffer containing 20 mM 2-mercaptoethanol. 2) Application of denatured TAG-1 to the column. 3) Washing step I: 3 CV of equilibration buffer. 4) Washing step II: 3 CV of solubilization buffer. 5) Elution with washing buffer II containing 0.25 M imidazole (Fluka).

C.3.2. Refolding and purification of renatured TAG-1_{Ig1-4}

Prior to refolding the concentration of denatured TAG-1_{Ig1-4} was adjusted to 4 mg/ml in elution buffer and DL-cysteine (Fluka) was added to a final concentration of 200 mM. TAG-1_{Ig1-4} was refolded using a quick dilution method into a buffer containing 0.8 M L-arginine (Fluka), 100 mM Tris-HCl pH 8.5, 320 mM NaCl. The final TAG-1_{Ig1-4} concentration during refolding was 40 mg/l. Dilution and refolding was carried out at 4 °C in 100 ml aliquots. The success of TAG-1_{Ig1-4} refolding was monitored by comparing the apparent molecular weight after SDS/PAGE (LAEMMLI, 1970) under reducing and non-reducing conditions.

Fully oxidized and refolded TAG-1_{Ig1-4} was concentrated 500 fold using an Amicon ultrafiltration unit with a YM10 membrane (Millipore) followed by a Vivaspin centrifugation filter (Vivascience GmbH, Germany), both with 10 kDa molecular weight cutoff. Concentrated TAG-1_{Ig1-4} was subjected to size exclusion chromatography with Superdex 200 (GE Healthcare) at a flow rate of 0.5 ml/cm² and a running buffer, containing 20 mM Tris-HCl pH 8.5, 150 mM NaCl. TAG-1 protein concentration in all experimental steps was determined as described (GILL AND VON HIPPEL, 1989).

C.3.3. Crystallization, data collection, phasing, and structure refinement

Human TAG-1_{Ig1-4} was crystallized using the sitting drop method by mixing 100 nl protein solution (4 mg/ml) and 100 nl crystallization buffer (12% polyethylene glycol 20000 (Fluka),

100 mM Tris-HCl pH 8.5, 200 mM KCl (Sigma)). Pipetting was done using an automatic liquid handling system (Cartesian Dispensing Systems, Genomic Solutions, USA). Needle shaped crystals were transferred into a drop containing 80 % (v/v) of the crystallization buffer supplemented with 20 % (v/v) ethylene glycol (Sigma) for five minutes prior to flash freezing in a 100 K cryostream (Oxford Cryosystems Ltd., UK). Diffraction data were collected at the X06SA beam line of the Swiss Light Source (Paul Scherrer Institut, Villigen, Switzerland). All diffraction images were processed using the program XDS (KABSCH, 1993). Phases were determined with the molecular replacement program MOLREP (VAGIN AND TEPLYAKOV, 1997) and chicken TAG-1 as search model (PDB code 1CS6). The solution obtained was improved by 30 cycles of rigid body refinement with REFMAC (COLLABORATIVE COMPUTATIONAL PROJECT NUMBER 4, 1994). The final model was obtained after several cycles of manual model building using COOT (EMSLEY AND COWTAN, 2004) and restrained refinement with REFMAC using separate anisotropic temperature factor tensors (TLS) (WINN *ET AL.*, 2001) for each Ig-domain.

C.3.4. Accession numbers

The coordinates and the structure factors of TAG-1_{Ig1-4} have been deposited in the Protein Data Bank (BERMAN *ET AL.*, 2000) under accession code 2AND.

C.3.5. Analysis of intermolecular interactions

The total buried surface value (B) of the TAG-1 dimer interfaces, the non-polar interface area ($f_{np}B$), the fraction of fully buried atoms (f_{bu}), and the residue propensity score (RP) were calculated by a web based server[#] (SAHA *ET AL.*, 2006). All parameter definitions are published (BAHADUR *ET AL.*, 2004). The shape complementarity value (S_C) was calculated by the program SC (LAWRENCE AND COLMAN, 1993) which is part of the CCP4 software package.

C.4. Results

C.4.1. Refolding and purification

Expression of TAG-1_{Ig1-4} in 1 l culture yielded 2 mg denatured and purified protein after the metal-affinity column. After initiation of refolding by dilution TAG-1_{Ig1-4} was checked daily for its apparent molecular mass by SDS-PAGE. Oxidation of all four disulfide bonds of TAG-1_{Ig1-4} was indicated by a shift from ~45 kDa to ~40 kDa due to the formation of four disulfide bonds. After seven days, the cysteines of were fully oxidized. By the gel filtration prior to crystallization, 0.1 mg monomeric, oxidized, purified TAG-1_{Ig1-4} per liter culture were obtained.

http://resources.boseinst.ernet.in/resources/bioinfo/interface

C.4.2. Crystallization, structure determination, and refinement

TAG-1_{Ig1-4} needle-like crystals grew within two weeks to a size of 5×5×70 μm^3 and belonged to space group $C222_1$ with one molecule per asymmetric unit. Data processing statistics is summarized in Table C.1.

Molecular replacement (MOLREP) with chicken TAG-1 (PDB code: 1CS6) as search model was successful to obtain initial phases. The final structure of TAG-1_{Ig1-4} had an overall root mean square (RMS) deviation of 1.52 Å[2] to the search model. Refinement with COOT and REFMAC yielded final R_{work}/R_{free} values of 23.5 %/27.9 %. After the refinement procedure reasonable geometrical values were obtained (see Table C.2).

Table C.1
Data Collection Statistics of TAG-1_{Ig1-4}

Data Set	TAG-1 native
Wavelength (Å)	1.07068
Resolution (Å)	20-3.1(3.26-3.07)[a]
Unique Reflections	7518 (950)
Completeness (%)	94.9 (76.7)
Average I/σ	12.9 (2.2)
R_{mrgd-f} (%)[b]	11.4 (59.7)
Spacegroup	$C222_1$
Cell Dimensions (Å)	a=46.7, b=106.8, c=161.3

[a] *Values in parentheses correspond to the highest resolution shell*

[b] *see (*DIEDERICHS AND KARPLUS, 1997*)*

C.4.3. U-shaped arrangement of the four N-terminal Ig-domains of TAG-1_{Ig1-4}

The four Ig-domains of human TAG-1_{Ig1-4} are arranged as a compact module (see Figure C.1), which allows strong interaction between the Ig1 and Ig4 domain as well as between Ig2 and Ig3 domain. The U-shaped overall structure was already predicted (RADER *ET AL.*, 1996) and reported (FREIGANG *ET AL.*, 2000) for the chicken ortholog. A similar U-shaped arrangement of four Ig-domains is shared also by the distantly related molecule hemolin from of the immune system of the giant silk moth *Hyalophora cecrophia* (SU *ET AL.*, 1998). Other cell adhesion molecules of the F11-family (F11, F3) and the L1-family (L1, NgCAM, and NrCAM), also have this extension. Compared to TAG-1 the NgCAM-related cell adhesion molecule NrCAM has additional 20 residues within the linker region, whereas the other molecules have no significant change in the length of the linker. A shorter linker than in TAG-1 probably would

not allow a bend after the Ig2 domain, which is required for the interaction between Ig2 and Ig3 domain in human and chicken TAG-1 and hemolin. In contrast, L1 may possess two different spatial arrangements of the four N-terminal Ig-domains, one with an U-shaped domain arrangement as reported here for TAG-1_{Ig1-4}, and an extended conformation, where the four N-terminal Ig-domains do not form a compact module (SCHÜRMANN *ET AL.*, 2001).

A comparison between human and chicken TAG-1_{Ig1-4}, and hemolin$_{Ig1-4}$ showed, that the root mean square (RMS) deviation of the domain pairs Ig1:Ig4 as well as Ig2:Ig3 is lower than the overall RMS deviation of the whole molecules. The RMS deviation between human and chicken TAG-$1_{Ig1:Ig4}$ is 0.82 Å; between human and chicken TAG-$1_{Ig2:Ig3}$ 0.75 Å; between the whole molecules TAG-1_{Ig1-4} of both organisms 1.52 Å. This indicates a tight interaction of the Ig-domains within the pairs Ig1:Ig4 and Ig2:Ig3 and the existence of two hinge joints, located between Ig1 and Ig2 and between Ig3 and Ig4 domain. A comparison of human TAG-1_{Ig1-4} with the chicken ortholog, using the program DYNDOM (HAYWARD AND LEE, 2002) located the residues of the hinge joints accordingly at the residues Leu^{102} and Gln^{296}. This suggests that the tertiary structure of TAG-1_{Ig1-4} can be seen as composed of two rigid groups: Ig1:Ig4 and Ig2:Ig3 domains. A similar comparison of human and chicken TAG-1_{Ig1-4} with hemolin$_{Ig1-4}$ gave further evidence for the two hinges.

Table C.2
Refinement Statistics of TAG-1_{Ig1-4}

Resolution (Å)	19.5 – 3.07	
Total no. of non-hydrogen atoms	2895	
Total no. of reflections	7090	
No. of reflections in test set	452	
R_{work}[a] (%)	23.5	
R_{free}[b] (%)	27.9	
r.m.s. distance from ideal geometry		
Bonds (Å)	0.007	
Angles (°)	1.078	
Ramachandran plot[c]		
Most favored regions	272	(86.1%)
Additionally allowed regions	37	(11.7%)
Generously allowed regions	3	(0.9%)
Disallowed regions	4	(1.3%)
No. non-glycine, non-proline, and non-end residues	316	(100%)
No. of glycine, proline and end residues	55	
Total no. of residues[d]	371	

Table C.2
Refinement Statistics of TAG-1_{Ig1-4}

[a] *R factor* = $\Sigma_{hkl}||F_{obs}| - k|F_{calc}||/\Sigma_{hkl}|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factors.
[b] For R_{free} the sum is extended over a subset of reflections, excluded from all stages of refinement.
[c] see (LASKOWSKI *ET AL.*, 1993)
[d] The residues 247-256 were excluded from refinement, because this region has no visible electron density.

C.4.4. An intermolecular β-strand pairing stabilizes the largest lattice contact between two TAG-1_{Ig1-4} molecules

Despite the structural similarity of human and chicken TAG-1_{Ig1-4}, the arrangement within the corresponding crystal lattice showed differences, which may be important to understand the molecular mechanism of homophilic interaction. The investigation of the lattice contacts of chicken TAG-1_{Ig1-4} resulted in a model for homophilic TAG-1 interaction, which was supported by two mutants of the chicken ortholog with mutations, located on the FG loop of the Ig2 domain. The deletion of the residues Ile^{187} to Ile^{190} as well as the combined exchange of His^{186} and Phe^{189} to alanine residues were sufficient to abolish myeloma cell-cell aggregation (FREIGANG *ET AL.*, 2000). In the human TAG-1_{Ig1-4} crystal structure, the lattice contact burying the largest surface (2240 Å^2) is composed of surface parts from Ig1 and Ig2

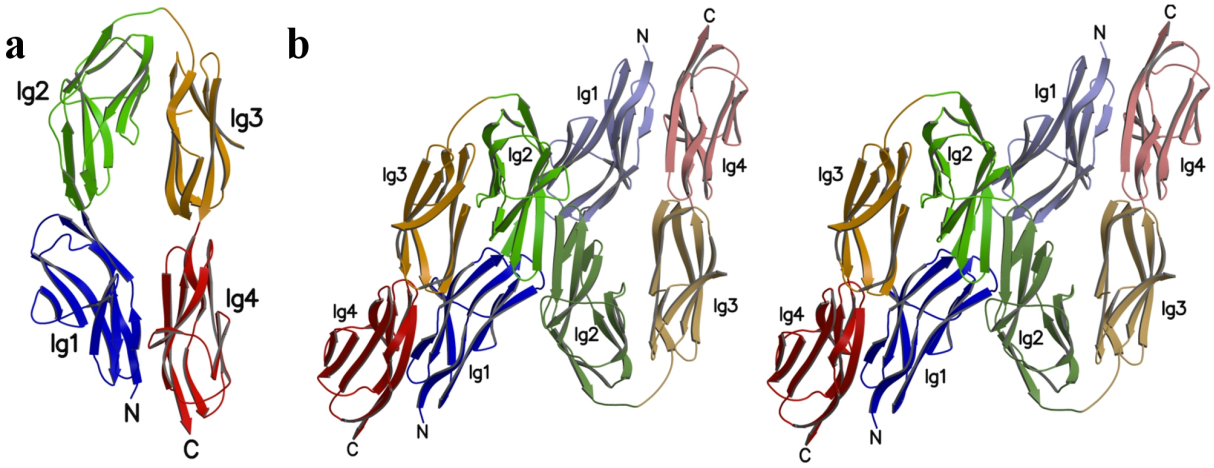


Figure C.1: Domain arrangement and groove to groove interaction of TAG-1_{Ig1-4}

a) Overall structure of human TAG-1_{Ig1-4}. The four N-terminal Ig-domains of human TAG-1 are arranged in an U-like manner. The extended linker between the Ig2 and Ig3 domain allows an apposition of the Ig3 and Ig4 domains, to the Ig1 and Ig2 domains. Specific interactions between Ig1 and Ig4, as well as between Ig2 and Ig3 stabilize the quaternary structure of the so called ligand binding module. The Ig pairs Ig1:Ig4 and Ig2:Ig3 form two rigid groups which are connected by a hinge, located at the residues Leu^{102} and Gln^{296}, which are found at the linkers between Ig1 to Ig2 and between Ig3 to Ig4. The residues 247-256 are not shown, because density was lacking.

b) Stereo view of the dimer of two TAG-1 ligand binding modules. The TAG-1_{Ig1-4} interaction site is located at the groove near the Ig1-Ig2 linker. Both molecules contact each other via their G strands (for details see Figure C.2a). Figure C.1 and C.2 were produced using the programs MOLSCRIPT (KRAULIS, 1991) and Raster3D (MERRITT AND BACON, 1997).

domain. Both molecules contact each other along the groove near the Ig1-Ig2 linker (see Fig C.1b). The FG loop of Ig2 domain is also part of this lattice contact which therefore has to be considered as a potential protein interaction site. All other lattice contacts buried surfaces, corresponding to average crystal lattice contacts (570 Å^2 for simple lattice contacts, see (JANIN, 1997); 1350 Å^2 for lattice contacts with two fold symmetry, see (JANIN AND RODIER, 1995)).

Deviant from the largest lattice contact found in the chicken ortholog, the two contacting human TAG-1$_{Ig1-4}$ molecules are related by two fold crystallographic symmetry. The rotation axis passes between adjacent G strands connected by six hydrogen bonds of the following backbone amino and carbonyl groups: Phe^{185}'O and Phe^{191}"N, Thr^{187}'N and Ser^{189}"O, Thr^{187}'O and Ser^{189}"N, Ser^{189}'N and Thr^{187}"O, Ser^{189}'O and Thr^{187}"N, and Phe^{191}'N and Phe^{185}"O. The apposition of both G strands thus results in an intramolecular antiparallel β-sheet (Figures C. 1b and C.2). The two residues, Phe^{185} and Thr^{187}, which are involved in the formation of the six hydrogen bonds of both participating β-strands, are in the elongated region of the FG loop, which is missing in the three other Ig-domains. In addition these central β-strands are framed by hydrophobic areas, which are formed mainly by the aromatic side chains of Phe^{143}, Phe^{191}, and Phe^{194} (see Figure C.2). These three residues stabilize each other and provide a hydrophobic pocket for the phenyl group of Phe^{185} of the corresponding dimer mate (see Figure C.2). The relevant distances of the phenylalanine side chains are in the range of 5.3-6.6 Å. This packing meets the characteristics of aromatic-aromatic interactions which are often found in the hydrophobic core of globular proteins (BURLEY AND PETSKO, 1985).

Another region contributing to the dimer interface is the segment from Glu^{21} to Glu^{24} from Ig1 which approaches the segment from Gln^{104} to Lys^{107} of Ig2 domain of the symmetry related molecule, but a detailed analysis of the interacting residues and its side chains is not possible because the side chain density does not allow further interpretation. Only an overall interlocking conformation of the participating TAG-1$_{Ig1-4}$ segments can be recognized.

The mode of TAG-1$_{Ig1-4}$ dimerization proposed in this work and the mode of oligomerization proposed for the chicken ortholog (FREIGANG *ET AL.*, 2000) both involve the same residues from β-strands F and G of the Ig2 domain. Each molecule contributes the same part of its surface to the homophilic dimer, because of the crystallographic twofold symmetry. However, the proposed oligomerization mechanism of the chicken ortholog is fundamentally different, because the interaction of two molecules involves different residues from both partners: The FG loop of the Ig2 domain of one molecule interacts with a protruding loop located between the β-strands C and E of the Ig3 domain from another molecule. This kind of interaction results in a string of interacting molecules, arranged in a zipper-like manner (see Figure C.3a for a schematic picture). The chicken TAG-1 zipper model requires that the interaction sites from adjacent molecules are different, whereas the dimer proposed here for

the human ortholog results in a pairwise interaction, involving identical interaction sites of each dimer mate.

As mentioned previously the residues of the FG loop from the Ig2 domain are important in both models. A clear difference between the chicken and human TAG-1_{Ig1-4} structure is the stabilization of the FG loop by the residue His^{182}. In human TAG-1_{Ig1-4} His^{182} is hydrogen bonded to the side chain of Ser^{135} (see Figure C.2), whereas in the chicken ortholog, the corresponding residue is hydrogen boded to the carbonyl group of Glu^{188}. Although the histidine side chain in both models differ by a rotation of almost 180 degrees, both FG loops have a very similar conformation.

C.4.5. A detailed analysis of human and chicken TAG-1_{Ig1-4} crystal lattices classifies the largest lattice contact of the human ortholog as a protein interaction site

In human TAG-1_{Ig1-4}, the largest surface of a lattice contact is 2240 $Å^2$, which is almost twice as large as the largest one of the chicken counterpart (1271 $Å^2$). Buried surfaces of lattice contacts and of protein interaction sites show large variations, but an area above 1200 $Å^2$ was proposed as characteristic of specific protein interaction sites (JANIN, 1997). Because buried surface areas of protein interaction sites and lattice contacts overlap, further criteria must be used to distinguish between both. S_C, $f_{np}B$, f_{bu}, and RP were calculated here. The combination of $f_{np}B$, f_{bu} and RP was found to allow discrimination between especially large lattice contacts (larger than 800 $Å^2$) and protein interaction sites (BAHADUR *ET AL.*, 2004). The shape complementarity value (S_C) (LAWRENCE AND COLMAN, 1993) is used to check how good two adjacent surfaces fit together by taking into account distance and angle between both surfaces. The calculated S_C-values were 0.70 for the human TAG-1_{Ig1-4} homophilic dimer and 0.63 for the zipper model of the chicken ortholog. A comparison with known protein interaction sites (Sc=0.70-0.78) and antibody-antigen complexes (Sc=0.62-0.68) (LAWRENCE AND COLMAN, 1993) shows that the S_C-value of the human TAG-1_{Ig1-4} homophilic dimer is in the range of known protein interaction sites, whereas in the chicken ortholog it is in the range of antibody-antigen complexes. Using the three values $f_{np}B$, f_{bu}, and RP, which do not correlate with the size (B) of the interface, as well as the combination of these values according to (BAHADUR *ET AL.*, 2004) classifies the dimer interface of the human ortholog as a protein interaction site, but the chicken TAG-1_{Ig1-4} zipper interface as a lattice contact. For human TAG-1_{Ig1-4}, the following values were determined: $f_{np}B$=1282, f_{bu}=0.66, and RP=3.14; and for the chicken counterpart: $f_{np}B$=740, f_{bu}=0.61, and RP=-0.91. According to (BAHADUR *ET AL.*, 2004), the combination of $f_{np}B$ and f_{bu} were sufficient to classify the chicken TAG-1_{Ig1-4} zipper interface

as a lattice contact. In the case of human TAG-1 the combination $f_{np}B$, f_{bu} and the large RP value (RP=3.14) classified the observed dimer interface as a protein interaction site.

C.5. Discussion

C.5.1. The binding module of TAG-1 is composed of two rigid groups

The U-shaped arrangement of the four N-terminal Ig-domains of human TAG-1 was also found in the chicken ortholog and moth hemolin with minor differences in the relative orientation of the single Ig-domains. The comparison of the RMS deviations showed, that the binding module is composed of two rigid groups, each formed by two Ig-domains: Ig1:Ig4 and Ig2:Ig3, which are connected by two hinges between Ig1 and Ig2 and between Ig3 and Ig4. The crystallographic TAG-$1_{Ig1\text{-}4}$ dimer with the largest interface area is formed by interaction of molecules, which interact along the groove formed by the Ig1-Ig2 linker. The existence of two rigid groups within the binding module may reflect the importance of a correct hinge angle as a prerequisite for TAG-1 dimerization. Two hinge joints as found here between both rigid groups of the binding module restrict the movement of both groups to a flapping movement around the hinge axis.

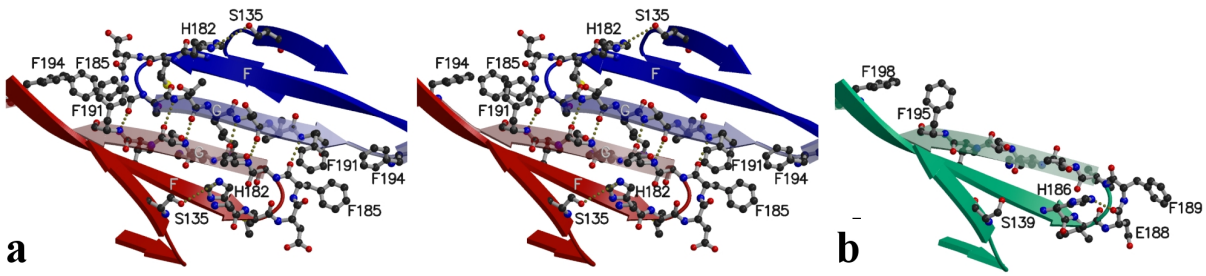


Figure C.2: The largest TAG-1 lattice contact

a) Stereo view: The antiparallel b-sheet of the crystallographic human TAG-$1_{Ig1\text{-}4}$ dimer is formed by apposition of two Ig2 G strands of TAG-$1_{Ig1\text{-}4}$ symmetry mates . The six intermolecular hydrogen bonds, which contribute to the stability of the dimer are indicated by dotted lines. On both ends of the contacting b-strands G^{blue} and G^{red} are hydrophobic pockets, each formed by Phe^{185}, Phe^{191} and Phe^{194}. On the bottom of the red molecule two hydrogen bonded residues, Ser^{135} and His^{182}, are shown. Phe^{143} which is part of the hydrophobic patch is not shown.

b) Detailed view of the Ig2 FG residues of chicken TAG-1 for comparison in the same orientation as the red molecule of Figure C.2a. Mutations which abolished myeloma cell-cell aggregation were targeted here: His^{186} and Phe^{189} point mutations to alanine as well a deletion mutant, lacking the four residues Ile^{187} to Ile^{190} (FREIGANG ET AL., 2000).

C.5.2. The largest lattice contact of TAG-1_{Ig1-4} may help to understand the molecular basis of homophilic interaction

As mentioned in a previous section, human and chicken TAG-1 are not only similar with respect to the overall structure of the four N-terminal Ig-domains, but also with respect to biochemical behavior in cellular assays, where TAG-1 as well as its binding partners can be replaced by its corresponding orthologs and vice versa (De Angelis *et al.*, 1999; Milev *et al.*, 1996; Pavlou *et al.*, 2002; Tsiotra *et al.*, 1993). This ability to substitute different molecules within the same experiment must reflect structural similarity and suggests that the mechanism for homophilic interaction of human TAG-1 as well as its chicken and rodent counterparts are similar. The zipper model proposed by (Freigang *et al.*, 2000), and the dimer model proposed here for homophilic interaction of the binding module are different. An important region for aggregation of myeloma cells by homophilic interaction of chicken TAG-1 was identified at the FG loop of Ig2 domain (Freigang *et al.*, 2000), which is involved in the largest lattice contact of both, human and chicken TAG-1_{Ig1-4} crystals. In the case of the human ortholog, the dimer with the largest buried surface is formed by intermolecular β-strand pairing of the two Ig2 G strands from different molecules. An intermolecular β-sheet is formed by both dimer mates, which are related by twofold crystallographic symmetry. However, in the chicken ortholog, the zipper formation is due to interaction of the Ig2 FG loop with the extended Ig3 CE loop, and contacting molecules are related by a crystallographic two fold screw axis.

Because both interaction modes seem to be mutually exclusive, but in accordance with the mutagenesis experiments of (Freigang *et al.*, 2000), they were compared carefully. Parameters were used, which can help to discriminate between lattice contacts and specific protein interaction sites. The interface area B and the shape complementarity value S_C are both larger for the interface of human TAG-1_{Ig1-4}. In antigen antibody complexes both parameters have values, which lie between lattice contacts and protein interaction sites (B=1500 ± 250 Å^2, see (Janin, 1997); S_C=0.64-0.68, see (Lawrence and Colman, 1993)). The B and S_C values of human TAG-1_{Ig1-4} are larger and the values for the chicken ortholog are smaller than the values for antibody-antigen complexes. Both values thus favor the model derived from the human TAG-1 structure over the zipper model, derived from the chicken ortholog. The non-polar interface area ($f_{np}B$), the fraction of fully buried atoms (f_{bu}), and the residue propensity score (RP) (Bahadur *et al.*, 2004) gave further support for the human TAG-1 derived dimer model, because the human TAG-1_{Ig1-4} dimer interface was classified as a specific protein interaction, whereas the zipper interface of the chicken ortholog was classified as a lattice contact. Conclusions about the *in vivo* situation can be wrong, because the classification procedure was correct for a distinct fraction of tested interfaces only. Statistically 7 % of validated protein interaction sites are classified as lattice contacts (false negatives) using the

$f_{np}B/f_{bu}$/RP classification. The same fraction (7 %) of false negatives was found for validated crystal lattice contacts with two fold symmetry, wrongly classified as protein interaction sites (BAHADUR *ET AL.*, 2004).

A closer view on those residues in the human ortholog, which correspond to the mutagenesis experiments of (FREIGANG *ET AL.*, 2000) shows that both combinations of mutations would probably destroy the central β-strand-pairing of the TAG-1_{Ig1-4} dimer. Firstly a deletion of the residues corresponding to chicken Ile^{187}-Ile^{190} in human TAG-1 would eliminate the dimerizing part if the Ig2 G strand by shortening the FG β-loop. In addition, the hydrophobic interaction of Phe^{185} with the hydrophobic pocket of the dimer mate would be abolished too. Secondly, a double alanine mutation of His^{182} and Phe^{185} would also abolish the hydrophobic interaction of Phe^{185} with the hydrophobic pocket or disturb the geometry of the whole FG β-loop, because of a stabilizing effect of the hydrogen bond of His^{182} to Ser^{135} in the human ortholog. Although the corresponding His^{186} in chicken TAG-1_{Ig1-4} does not form a hydrogen bond with the corresponding Ser^{139}, but with the main chain carbonyl of Glu^{188} (see Figure C.2b) the conformation of the FG loop is similar to that observed in the human TAG-1_{Ig1-4} structure. Since there are no single alanine mutants, it is unknown, if either the histidine or the phenylalanine residue are important for homophilic interaction alone, or both together. The analysis of the largest lattice contact as well as the detailed analysis of the Ig2 FG loop give strong hints, that this Ig2 FG loop is involved in a homophilic interaction mechanism.

C.5.3. The TAG-1 protein interaction site is in accordance with other experiments

We propose, that the characteristic of TAG-1 dimerization is a groove to groove interaction of the Ig1/Ig2 domains, which requires a correct angular orientation of the rigid groups Ig1:Ig4 and Ig2:Ig3 with respect to the hinge axis, and that TAG-1 dimer stability results from the hydrophobic interaction of the Phe^{185} side chain with the hydrophobic pocket of the dimer mate and from intermolecular β-strand pairing of the Ig2 G strands. Therefore, the proposed dimerization of TAG-1 may involve an induced fit mechanism (KOSHLAND, 1958), requiring domain flexibility between Ig1 and Ig2 and small side chain readjustments within the FG loop of the Ig2 domain during dimer formation.

The observed crystallographic dimer of two human TAG-1_{Ig1-4} molecules thus may promote homophilic *trans* interaction *in vivo*. Homophilic cell-cell adhesion mediated by TAG-1 dimers contradict a zipper model as proposed for the chicken ortholog, because in human TAG-1 both dimer mates interact with identical surface sites, namely the G strands located at the Ig1-Ig2 groove, whereas the proposed zipper units of the chicken ortholog

contact each other via different surface sites, the Ig2 FG loop of one molecule and the extended Ig3 CE loop of a second TAG-1 molecule. Another zipper based mechanism was proposed for cell-cell adhesion mediated by cadherins, which involves two fold crystallographic symmetry between all molecular *cis* and *trans* contacts of the proposed zipper. The cadherin model requires two different interaction sites on each molecule, which interact only with the respecting site of the interaction partner (SHAPIRO *ET AL.*, 1995).

In the case of human TAG-1, S2 cell-cell aggregation via the FnIII-domains was reported (TSIOTRA *ET AL.*, 1996). Antibody mapping experiments, in combination with myeloma cell-cell interaction studies had narrowed down the importance of the FnIII-domains to the fourth domain, for TAG-1 induced cell-cell interaction (KUNZ *ET AL.*, 2002). These findings led to an extended zipper model, where *trans*-interacting TAG-1 molecules forming a zipper, interact in addition, with their FnIII-domains in *cis*, by dimer formation via the fourth FnIII-domain (see Figure 5, reference (KUNZ *ET AL.*, 2002)). There are no experiments which show that the *cis* interaction of TAG-1 needs a preformed "*trans* zipper", or that TAG-1 molecules forming

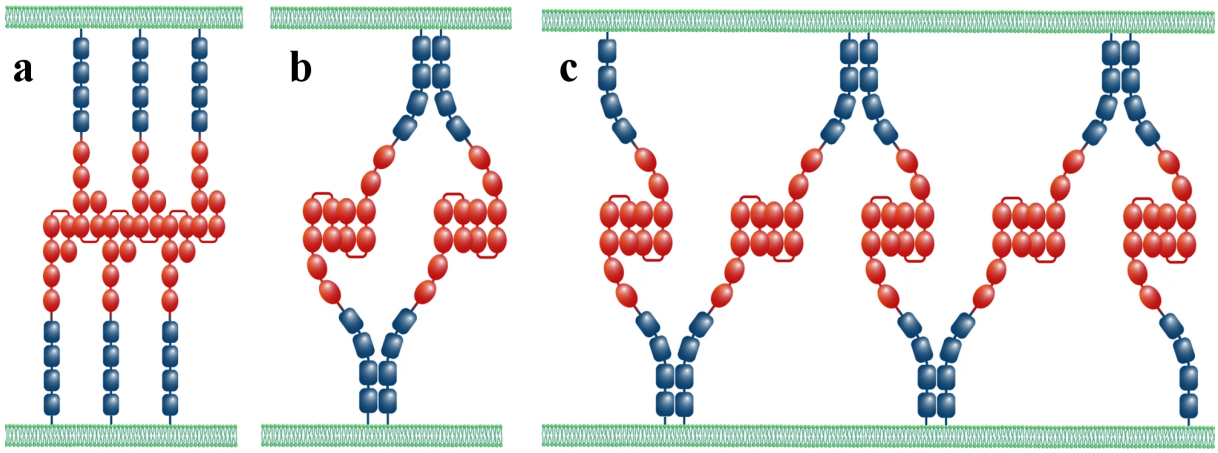


Figure C.3: Comparison of putative TAG-1 models

N-terminal Ig-domains are indicated in red, C-terminal FnIII-domains in blue. All TAG-1 molecules are GPI-anchored in the cell membrane, by the fourth FnIII-domain.

(a) Zipper model, as proposed for chicken TAG-1_{Ig1-4} (FREIGANG ET AL., 2000). All molecules have two different protein interaction sites, which are located on the Ig2 and Ig3 domain respectively.

(b) The four molecule model is the smallest possible trans-cis complex in accordance with the crystallographic dimer of human TAG-1_{Ig1-4}, where all molecules are involved in trans and cis interactions simultaneously.

(c) The multiple molecule model is a variation of (b), where many molecules participate in a linear arrangement. The flexibility between cis and trans binding sites allows a curvilinear arrangement of consecutive TAG-1 dimers, which is not implied by the schematic view.

a homodimer but not a zipper can not interact in *cis* via its FnIII-domains with other TAG-1 molecules, and nothing is known, whether TAG-1 molecules interact in *cis* without interacting simultaneously in *trans*. It is known however, that blocking of TAG-1 mediated myeloma cell-cell aggregation is possible using monoclonal antibodies against the ligand binding module, namely the four N-terminal Ig-domains, or the fourth FnIII-domain (KUNZ *ET AL.*, 2002). This suggests that homophilic cell-cell aggregation by TAG-1 can only occur, if at least some of the involved TAG-1 molecules bind other TAG-1 molecules located in *cis* and *trans* simultaneously.

C.5.4. The four molecule model and the multiple molecule model

On the basis of the proposed interaction site of human TAG-1, two different models of cell-cell interaction can be proposed. The *four molecule model*: a model with only four TAG-1 molecules (see Figure C.3b), and the *multiple molecule model*, a model with an unbranched curvilinear string of participating TAG-1 molecules (see Figure C.3c). The interacting sites are similar for both models and comply with the requirements of simultaneous *cis* interaction via the fourth FnIII-domain and *trans* interaction via dimerization of two binding modules (KUNZ *ET AL.*, 2002).

The *four molecule model* represents the model with the smallest possible number of participating TAG-1 molecules involving simultaneous *cis* and *trans* interactions of each participating molecule (see Figure C.3b). In the *four molecule model*, two molecules from the same membrane are connected to a pair of *cis* interacting FnIII-domains. In addition each molecule is connected in *trans* independently with a TAG-1 molecule originating from a second *cis* dimer of the opposing membrane.

The *multiple molecule model* (see Figure C.3c) is a slightly modified version of the *four molecule model*, where each TAG-1 molecule of the *cis* dimer is connected in *trans* with a TAG-1 molecule, originating from two different *cis* dimers of the opposing membrane. It slightly resembles the already proposed zipper-like model (compare with Figure C.3a and Reference (KUNZ *ET AL.*, 2002)), but does not allow for two dimensional network of TAG-1 molecules at the contact interface of two interacting cells.

The formation of *trans-cis* complexes of the *multiple molecule* and the *four molecule* type can occur stepwise. In a first aggregation event the *cis* interaction of the two FnIII-domains can be established followed by *trans* association of the binding modules. It was observed that chicken TAG-1 adopts a back folded conformation, where the four N-terminal Ig-domains are located towards the membrane in close proximity to the FnIII-domains (RADER *ET AL.*, 1993). Probably back folded TAG-1 molecules can interact with already *trans* dimerized molecules with their fourth FnIII, which could induce the release of the ligand binding module from the

back folded orientation onto an orientation towards the subsequent *trans* interaction partner, a TAG-1 molecule presented by a neighboring cell.

Without further experimental evidence it is not possible to verify one of the possible interaction models, the *four molecule model*, the *multiple molecule model*, or the zipper model, and reject the remaining ones.

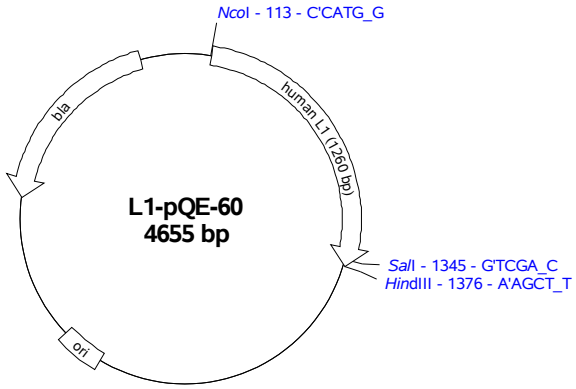
Appendix

Expression vector L1-pQE-60

Expression vector L1-pQE-60 was cloned as described (**MÖRTL, 2000**).

Vector map of L1-pQE-60:

L1 (Ig1-4)	*115-1375*
bla (resistance)	*4450-3590*
ori	*2832*



DNA Sequence of L1 insert:

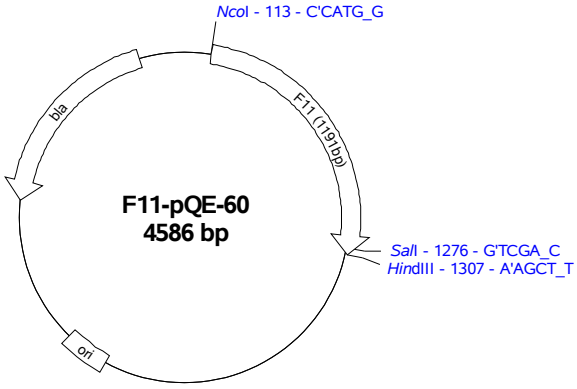
```
 101 AGGAGAAATT AACCATGGCT ATCCAGATCC CCGAGGAATA TGAAGGACAC CATGTGATGG AGCCACCTGT CATCACGGAA CAGTCTCCAC GGCGCCTGGT
 201 TGTCTTCCCC ACAGATGACA TCAGCCTCAA GTGTGAGGCC AGTGGCAAGC CCGAAGTGCA GTTCCGCTGG ACGAGGGATG GTGTCCACTT CAAACCCAAG
 301 GAAGAGCTGG GTGTGACCGT GTACCAGTCG CCCCACTCTG GCTCCTTCAC CATCACGGGC AACAACAGCA ACTTTGCTCA GAGGTTCCAG GGCATCTACC
 401 GCTGCTTTGC CAGCAATAAG CTGGGCACCG CCATGTCCCA TGAGATCCGG CTCATGGCCG AGGGTGCCCC CAAGTGGCCA AAGGAGACAG TGAAGCCCGT
 501 GGAGGTGGAG GAAGGGGAGT CAGTGGTTCT GCCTTGCAAC CCTCCCCCAA GTGCAGAGCC TCTCCGGATC TACTGGATGA ACAGCAAGAT CTTGCACATC
 601 AAGCAGGACG AGCGGGTGAC GATGGGCCAG AACGGCAACC TCTACTTTGC CAATGTGCTC ACCTCCGACA ACCACTCAGA CTACATCTGC CACGCCCACT
 701 TCCCAGGCAT TAGGACCATC ATTCAGAAGG AACCCATTGA CCTCCGGGTC AAGGCCACCA ACAGCATGAT TGACAGGAAG CCGCGCCTGC TCTTCCCCAC
 801 CAACTCCAGC ACGCACCTGG TGGCCTTGCA GGGGCAGCCA TTGGTCCTGG AGTGCATCGC CGAGGGCTTT CCCACGCCCA CCATCAAATC GGTGCGCCCC
 901 AGTGGCCCCA TGCCAGCCGA CCGTGTCACC TACCAGAACC ACAACAAGAC CCTGCAGCTG CTGAAAGTGG GCGAGGAGGA TGATGGCGAG TACCGCTGCC
1001 TGGCCGAGAA CTCACTGGGC AGTGCCCGGC ATGCGTACTA TGTCACCGTG GAGGCTGCCC CGTACTGGCT GCACAAGCCC CAGAGCCATC TATATGGGCC
1101 AGGAGAGACT GCCCGCCTGG ACTGCCAAGT CGAGGGCAGG CCCCAACCAG AGGTCACCTG GAGAATCAAC GGGATCCCTG TGGAGGAGCT GGCCAAAGAC
1201 CAGAAGTACC GGATTCAGCG TGGCGCCCTG ATCCTGAGCA ACGTGCAGCC CAGTGACACA ATGGTGACCC AATGTGAGGC CCGCAACCGG CACGGGCTCT
1301 TGCTGGCCAA TGCCTACATC TACGTTGTCC AGCTGCCAGC CAAGGTCGAC AGATCTCATC ACCATCACCA TCACTAAGCT T
```

Expression vector F11-pQE-60

Expression vector F11-pQE-60 was cloned as described (**MÖRTL, 2000**).

Vector map of F11-pQE-60:

F11 (Ig1-4)	*115-1306*
bla (resistance)	*4381-3521*
ori	*2763*



DNA Sequence of F11 insert:

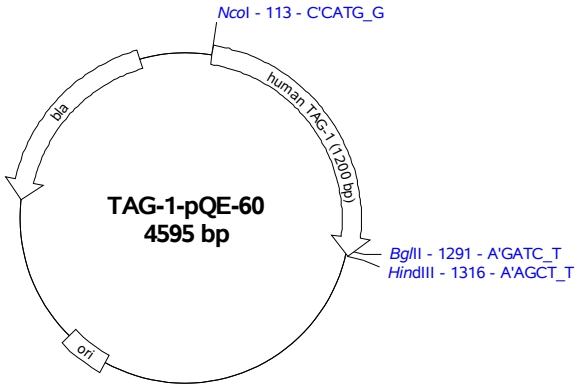
```
 101 AGGAGAAATT AACCATGGCT ACCCATTTTT CAGAGGAAGG AAACAAAGGT TATGGACCCG TTTTTGAGGA GCAACCCATT GATACTATCT ATCCAGAAGA
 201 ATCTTCAGAT GGACAAGTTT CAATGAACTG CAGGGCTCGA GCAGTTCCTT TTCCTACTTA CAAGTGGAAA CTCAACAACT GGGATATTGA TTTAACAAAG
 301 GATCGCTACA GTATGGTAGG AGGCAGACTT GTTATCAGTA ATCCAGAGAA ATCAAGGGAT GCTGGAAAAT ATGTCTGTGT AGTATCGAAT ATCTTTGGAA
 401 CTGTCAGAAG CAGTGAAGCC ACTCTGAGTT TTGGATATCT GGATCCTTTT CCACCAGAGG AACACTATGA GGTTAAAGTA CGAGAAGGTG TTGGAGCAGT
 501 GCTTCTTTGT GAGCCTCCTT ACCACTACCC AGATGATCTC AGTTATCGCT GGCTTCTAAA TGAGTTCCCA GTATTCATTG CTTTGGACAG ACGACGGTTT
 601 GTGTCTCAGA CGAATGGCAA TCTCTACATT GCAAATGTAG AAGCATCAGA TAAGGGAAAT TATTCATGTT TTGTGTCCAG CCCTTCGATA ACAAAGAGTG
 701 TATTCAGTAA ATTTATTCCA CTTATTCCAC AGGCCGATCG TGCCAAAGTG TATCCTGCTG ATATCAAGGT GAAGTTCAAG GACACGTATG CTCTCCTGGG
 801 GCAAAATGTG ACACTGGAGT GTTTTGCTCT TGGGAATCCG GTTCCTGAAC TCAGATGGAG TAAATACCTT GAACCTATGC CAGCCACTGC TGAGATAAGC
 901 ATGTCTGGTG CAGTTCTTAA GATCTTTAAT ATTCAGTATG AAGATGAAGG ACTTTATGAG TGTGAAGCTG AAAACTATAA AGGAAAAGAT AAACATCAAG
1001 CAAGAGTTTA TGTGCAAGCC TCTCCAGAGT GGGTAGAACA TATCAACGAC ACAGAGAAGG ATATAGGCAG CGATCTCTAC TGGCCGTGTG TAGCTACAGG
1101 CAAGCCTATT CCAACAATTA GGTGGTTGAA GAACGGTGTC TCGTTCCGGA AAGGTGAACT ACGAATTCAA GGCCTGACCT TTGAAGATGC TGGCATGTAT
1201 CAGTGTATAG CAGAAAACGC ACATGGAATT ATTTATGCAA ATGCTGAACT GAAGATTGTG GCTTCACCTC CTACTGTCGA CAGATCTCAT CACCATCACC
1301 ATCACTAAGC TTAATTAGCT GAGCTTGGAC TCCTGTTGAT AGATCCAGTA
```

Expression vector TAG-1-pQE-60

Expression vector TAG-1-pQE-60 was cloned as described (**MÖRTL, 2000**).

Vector map of TAG-1-pQE-60:

human TAG-1 (Ig1-4)	*115-1315*
bla (resistance)	*4390-3530*
ori	*2772*



DNA Sequence of human TAG-1 insert:

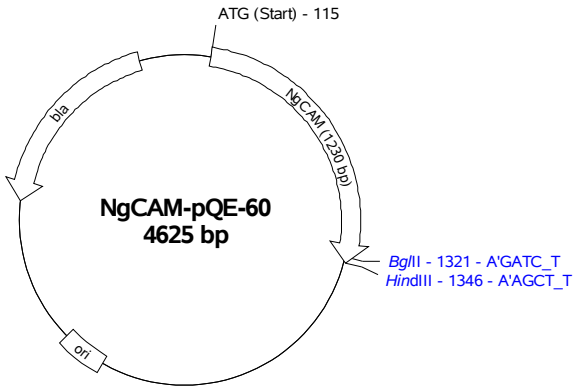
```
 101 AGGAGAAATT AACCATGGCC CTGGGATCCC AAACCACCTT CGGGCCTGTC TTTGAAGACC AGCCCCTCAG TGTGCTATTC CCAGAGGAGT CCACGGAGGA
 201 GCAGGTGTTG CTGGCATGCC GCGCCCGGGC CAGCCCTCCA GCCACCTATC GGTGGAAGAT GAATGGTACC GAGATGAAGC TGGAGCCAGG TTCCCGTCAC
 301 CAGCTGGTGG GGGGCAACCT GGTCATCATG AACCCCACCA AGGCACAGGA TGCCGGGGTC TACCAGTGCC TGGCCTCCAA CCCAGTGGGC ACCGTTGTCA
 401 GCAGGGAGGC CATCCTCCGC TTCGGCTTTC TGCAGGAATT CTCCAAGGAG GAGCGAGACC CAGTGAAAGC TCATGAAGGC TGGGGGGTGA TGTTGCCCTG
 501 TAACCCACCT GCCCACTACC CAGGCTTGTC CTACCGCTGG CTCCTCAACG AGTTCCCCAA CTTCATCCCG ACGGACGGGC GTCACTTCGT GTCCCAGACC
 601 ACAGGGAACC TGTACATTGC CCGAACCAAT GCCTCAGACC TGGGCAACTA CTCCTGTTTG GCCACCAGCC ACATGGACTT CTCCACCAAG AGCGTCTTCA
 701 GCAAGTTTGC TCAGCTCAAC CTGGCTGCTG AAGATACCCG GCTCTTTGCA CCCAGCATCA AGGCCCGGTT CCCAGCAGAG ACCTATGCAC TGGTGGGGCA
 801 GCAGGTCACC CTGGAGTGCT TCGCCTTTGG GAACCCTGTC CCCCGGATCA AGTGGCGCAA AGTGGACGGC TCCCTGTCCC CGCAGTGGAC CACAGCTGAG
 901 CCCACCCTGC AGATCCCCAG CGTCAGCTTT GAGGATGAGG GCACCTACGA GTGTGAGGCG GAGAACTCCA AGGGCCGAGA CACCGTGCAG GGCCGCATCA
1001 TCGTGCAGGC TCAGCCTGAG TGGCTAAAAG TGATCTCGGA CACAGAGGCT GACATTGGCT CCAACCTGCG TTGGGGCTGT GCAGCCGCCG GCAAGCCCCG
1101 GCCTACAGTG CGCTGGCTGC GGAACGGGGA GCCTCTGGCC TCCCAGAACC GGGTGGAGGT GTTGGCTGGG GACCTGCGGT TCTCCAAGCT GAGCCTGGAA
1201 GACTCGGGCA TGTACCAGTG TGTGGCAGAG AATAAGCACG GTACCATCTA CGCCAGCGCC GAGCTAGCCG TGCAAGCACT CGCCCCTGAC AGATCTCATC
1301 ACCATCACCA TCACTAAGCT TA
```

Expression vector NgCAM-pQE-60

Expression vector NgCAM-pQE-60 was cloned as described (**Mörtl, 2000**).

Vector map of NgCAM-pQE-60:

NgCAM (Ig1-4)	*115-1345*
bla (resistance)	*4420-3560*
ori	*2802*



DNA Sequence of NgCAM insert:

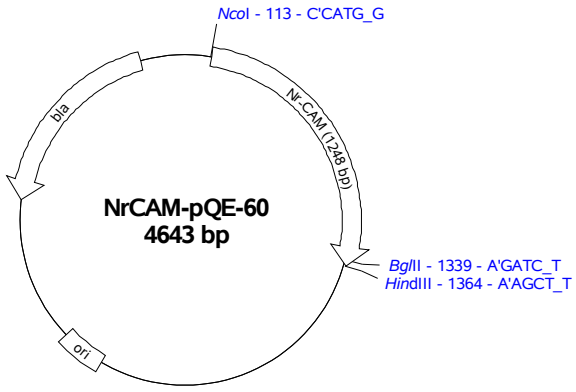
```
 101 AGGAGAAATT AAATGAATC ACCATTCCCC CGGAGTATGG TGCACACGAT TTCCTGCAGC CCCCCGAGCT GACGGAGGAA CCCCCGGAAC AACTCGTGGT
 201 CTTCCCCAGT GATGACATCG TCCTCAAATG CGTGGCCACC GGGAACCCCC CCGTCCAGTA CCGATGGAGC CGTGAGGATC AGCCCTTCGT CCCCGAGGAG
 301 CACGGGGGGG TCTCGGTGGT CCCCGGATCG GGGACTTTGG TCATCAACGC CACGTTGGCC GCGCGGCTCC AGGGGCGCTT CCGCTGCTTC GCCACCAACG
 401 CGTTGGGCAC CGCTGTGTCT CCCGAGGCCA ACGTCATCGC CGAGAACACT CCGCAGTGGC CGGAGAAGAA GGTGACCCCG GTGGAGGTGG AGGAGGGGGA
 501 CCCCGTGGTG CTGCCCTGTG ACCCCCCCGA GAGCGCTGTG CCCCCTAAAA TCTATTGGCT CAACAGCGAC ATCGTTCACA TCGCTCAGGA CGAGAGGGTA
 601 TCGATGGGGC AGGATGGGAA CCTCTACTTC TCCAACGCCA TGGTGGGCGA CAGCCACCCC GACTACATCT GCCACGCTCA CTTCCTCGGC CCCCGCACCA
 701 TCATCCAGAA GGAGCCCCTC GACCTCCGCG TGGCCCCCAG TAATGCGGTT CGGTCCCGCC GCCCCCGCCT GCTGCTGCCC CGCGACCCCC AAACGACCAC
 801 CATCGCCCTC CGGGGGGGCA GCGTCGTGTT GGAGTGCATC GCTGAGGGGC TCCCCACTCC ATGGGTCCGA TGGCGGCGGC TGAACGGCCC CCTCCCCACC
 901 CGGGCGGCGT TGGAGAACTT CAACAAAACG CTGCGGCTGT GGGGGGTGAC GGAGAGCGAC GACGGGGAGT ACGAATGTGT GGCTGAGAAC GGGAGGGGGA
1001 CGGCCAGGGG CACCCACAGC GTCACCGTGG AGGCCGCCCC ATATTGGGTG CGGCGGCCAC AGAGTGGGGT CTTCGGGCCG GGGGAGACGG CGAGGCTGGA
1101 CTGCGAGGTG GGGGGGAAAC CCCGACCCCA AATCCAATGG AGCATCAATG GGGTCCCCAT CGACGCGCTG CCGGGCGCGG AGCGGCGGTG GCTGCGGGGC
1201 GGCGCTTTGG TGCTTCCGGA GCTGCGGCCG AACGACAGCG CGGTGCTGCA GTGCGAGGCG AGGAACCGCC ACGGCCGCCT ATTGGCCAAC GCCTTCCTGC
1301 ACGTCGTGGA GCTGCCGGTC AGATCTCATC ACCATCACCA TCACTAAGCT T
```

Expression vector NrCAM-pQE-60

Expression vector NrCAM-pQE-60 was cloned as described (**Mörtl, 2000**).

Vector map of NrCAM-pQE-60:

NrCAM (Ig1-4)	*115-1363*
bla (resistance)	*4438-3578*
ori	*2820*



DNA Sequence of NrCAM insert:

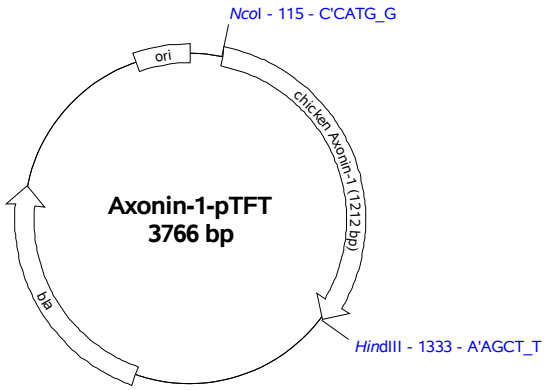
```
  101 AGGAGAAATT AACCATGGCA TTGGATGTAC CTCTTGATTC AAAACTTCTA GAAGAATTGT CTCAACCTCC AACAATAACT CAGCAGTCTC CAAAAGATTA
  201 CATTGTTGAC CCTCGAGAGA ATATTGTAAT ACAATGTGAA GCAAAAGGAA AACCACCTCC TAGCTTCTCC TGGACGCGCA ATGGAACTCA TTTTGATATA
  301 GATAAAGATG CACAGGTAAC AATGAAACCA AATTCAGGAA CCCTTGTTGT AAATATTATG AATGGTGTGA AGGCAGAAGC ATATGAAGGA GTATACCAGT
  401 GTACAGCAAG GAATGAAAGA GGAGCAGCCA TTTCCAACAA TATTGTTATA CGGCCATCTA GATCCCCTTT GTGGACTAAA GAAAAACTAG AACCAAATCA
  501 TGTTCGAGAA GGTGATTCCC TAGTACTAAA CTGCAGACCT CCTGTTGGCT TACCACCACC TATAATATTT TGGATGGATA ATGCTTTCCA AAGGCTGCCT
  601 CAAAGTGAAA GAGTTTCTCA AGGTCTCAAT GGAGACCTTT ATTTTTCTAA TGTACAACCA GAGGACACCC GTGTGGACTA TATCTGCTAC GCGAGATTTA
  701 ATCACACACA AACTATACAG CAGAAACAAC CCATTTCTGT AAAAGTCTTT TCAACCAAGC CAGTTACAGA AAGGCCACCA GTTCTTCTTA CACCAATGGG
  801 CAGCACAAGT AACAAAGTGG AACTGAGAGG AAATGTTCTT TTGTTGGAAT GCATCGCAGC AGGATTACCC ACACCAGTAA TCCGCTGGAT TAAAGAGGGT
  901 GGTGAACTGC CAGCCAACAG AACGTTTTTT GAAAATTTTA AGAAAACTCT CAAGATTATA GACGTCTCTG AAGCTGACTC TGGGAACTAC AAATGTACAG
 1001 CAAGAAATAC ATTGGGTTCT ACTCATCATG TCATTTCGGT AACTGTAAAA GCTGCCCCAT ACTGGATAAC AGCACCCAGG AACTTAGTAT TGTCTCCTGG
 1101 AGAAGATGGG ACATTGATCT GCAGAGCTAA TGGCAACCCA AAACCTAGCA TAAGCTGGTT AACAAATGGC GTTCCCATAG CAATTGCCCC AGAAGATCCT
 1201 AGCAGAAAGG TAGATGGGGA TACCATTATT TTCTCAGCTG TGCAAGAACG GTCAAGTGCT GTTTATCAGT GCAATGCTTC TAATGAGTAT GGATACTTGC
 1301 TGGCAAATGC ATTTGTGAAT GTTCTTGCTG AGCCACCAAG ATCTCATCAC CATCACCATC ACTAAGCT T
```

Expression vector Axonin-1-pTFT

Expression vector Axonin-1-pTFT served as source for all other pTFT constructs. The restriction sites used for cloning were the NcoI (*CCATGG*) and the HindIII (*AAGCTT*) site.

Vector map of Axonin-1-pTFT:

Axonin-1 (Ig1-4)	*117-1329*
bla (resistance)	*2087-2945*
ori	*3705*



DNA Sequence of vector Axonin-1-pTFT

```
   1 ACCCGACACC ATCGAAATTA ATACGACTCA CTATAGGGAG ACCACAACGG TTTCCCGAAT TGTGAGCGGA TAACAATAGA AATAATTTTG TTTAACTTTA
 101 AGAAGGAGAT ATATCCATGG CTCACCACCA CCACCACCAC ATCGAGGGAA GGCAAGCTCA GAGTGGCATG AGGAGCTACG GGCCGGTGTT TGAGGAGCAG
 201 CCAGCACACA CCCTGTTCCC TGAAGGCTCA GCAGAAGAGA AAGTGACGCT GACGTGTCGA GCCAGAGCCA ACCCCCCTGC AACATACAGA TGGAAGATGA
 301 ACGGCACAGA GCTGAAGATG GGGCCGGATT CACGGTACAG GTTGGTTGCT GGGGACCTGG TGATAAGCAA CCCAGTGAAA GCAAAGGATG CAGGCTCCTA
 401 CCAGTGTGTG GCAACCAACG CCCGTGGCAC AGTGGTCAGC AGGGAAGCAT CCCTCCGCTT TGGCTTTTTG CAGGAATTCT CTGCAGAAGA GCGAGACCCC
 501 GTGAAGATCA CCGAAGGTTG GGGAGTGATG TTCACCTGCA GCCCCCCTCC CCATTACCCA GCCTTGTCCT ACCGGTGGCT TCTGAATGAG TTTCCCAACT
 601 TCATCCCAGC CGACGGAAGG CGTTTCGTCT CTCAGACCAC AGGAAACCTC TACATTGCCA AGACAGAGGC TTCCGACCTC GGCAACTACT CGTGCTTTGC
 701 CACCAGCCAC ATCGACTTCA TCACCAAGAG CGTTTTCAGC AAGTTCTCCC AGCTCAGCCT GGCTGCCGAG GACGCCAGGC AGTATGCACC CAGCATAAAA
 801 GCCAAGTTCC CTGCAGACAC CTACGCTCTG ACCGGGCAGA TGGTGACTCT GGAGTGTTTC GCCTTTGGGA ACCCCGTTCC TCAAATAAAG TGGAGGAAGC
 901 TGGATGGCTC GCAGACCTCC AAGTGGCTCT CCAGTGAGCC CCTCCTGCAC ATCCAAAATG TTGACTTTGA AGATGAAGGC ACTTATGAGT GCGAGGCTGA
1001 AAACATCAAA GGGAGGGACA CCTACCAGGG CCGCATCATC ATTCACGCTC AGCCGGACTG GCTGGATGTG ATCACGGACA CGGAAGCTGA CATTGGGTCC
1101 GACCTGCGCT GGAGCTGCGT GGCATCCGGC AAACCCCGGC CAGCAGTGCG GTGGCTGCGG GATGGGCAAC CTCTGGCCTC CCAGAACCGC ATCGAAGTGA
1201 GCGGTGGAGA GCTGAGATTT TCCAAGCTAG TCCTGGAGGA CTCTGGCATG TATCAGTGTG TGGCTGAGAA CAAGCATGGC ACAGTATATG CAAGTGCTGA
1301 ATTAACAGTG CAAGCCTTAG CACCAGATTG ATAAGCTTCA GTCCCGGGCA GTGGATCCGG CTGCTAACAA AGCCCGAAAG GAAGCTGAGT TGGCTGCTGC
1401 CACCGCTGAG CAATAACTAG CATAACCCCT TGGGGCCTCT AAACGGGTCT TGAGGGGTTT TTTGCTGAAA GGAGGAACTA TATCCGGATC GAGATCCCCA
1501 CGCGCCCTGT AGCGGCGCAT TAAGCGCGGC GGGTGTGGTG GTTACGCGCA GCGTGACCGC TACACTTGCC AGCGCCCTAG CGCCCGCTCC TTTCGCTTTC
1601 TTCCCTTCCT TTCTCGCCAC GTTCGCCGGC TTTCCCCGTC AAGCTCTAAA TCGGGGCATC CCTTTAGGGT TCCGATTTAG TGCTTTACGG CACCTCGACC
1701 CCAAAAAACT TGATTAGGGT GATGGTTCAC GTAGTGGGCC ATCGCCCTGA TAGACGGTTT TTCGCCCTTT GACGTTGGAG TCCACGTTCT TTAATAGTGG
1801 ACTCTTGTTC CAAACTGGAA CAACACTCAA CCCTATCTCG GTCTATTCTT TTGATTTATA AGGGATTTTG CCGATTTCGG CCTATTGGTT AAAAAATGAG
1901 CTGATTTAAC AAAAATTTAA CGCGAATTTT AACAAAATAT TAACGTTTAC AATTTCAGGT GGCACTTTTC GGGGAAATGT GCGCGGAACC CCTATTTGTT
2001 TATTTTTCTA AATACATTCA AATATGTATC CGCTCATGAG ACAATAACCC TGATAAATGC TTCAATAATA TTGAAAAAGG AAGAGTATGA GTATTCAACA
2101 TTTCCGTGTC GCCCTTATTC CCTTTTTTGC GGCATTTTGC CTTCCTGTTT TTGCTCACCC AGAAACGCTG GTGAAAGTAA AAGATGCTGA AGATCAGTTG
2201 GGTGCACGAG TGGGTTACAT CGAACTGGAT CTCAACAGCG GTAAGATCCT TGAGAGTTTT CGCCCCGAAG AACGTTTTCC AATGATGAGC ACTTTTAAAG
2301 TTCTGCTATG TGGCGCGGTA TTATCCCGTA TTGACGCCGG GCAAGAGCAA CTCGGTCGCC GCATACACTA TTCTCAGAAT GACTTGGTTG AGTACTCACC
2401 AGTCACAGAA AAGCATCTTA CGGATGGCAT GACAGTAAGA GAATTATGCA GTGCTGCCAT AACCATGAGT GATAACACTG CGGCCAACTT ACTTCTGACA
```

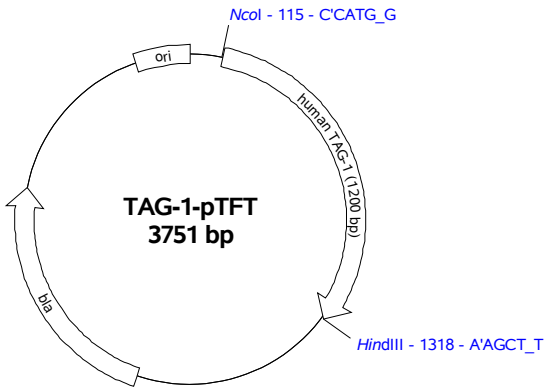
```
2501 ACGATCGGAG GACCGAAGGA GCTAACCGCT TTTTTGCACA ACATGGGGGA TCATGTAACT CGCCTTGATC GTTGGGAACC GGAGCTGAAT GAAGCCATAC
2601 CAAACGACGA GCGTGACACC ACGATGCCTG TAGCAATGGC AACAACGTTG CGCAAACTAT TAACTGGCGA ACTACTTACT CTAGCTTCCC GGCAACAATT
2701 AATAGACTGG ATGGAGGCGG ATAAAGTTGC AGGACCACTT CTGCGCTCGG CCCTTCCGGC TGGCTGGTTT ATTGCTGATA AATCTGGAGC CGGTGAGCGT
2801 GGGTCTCGCG GTATCATTGC AGCACTGGGG CCAGATGGTA AGCCCTCCCG TATCGTAGTT ATCTACACGA CGGGGAGTCA GGCAACTATG GATGAACGAA
2901 ATAGACAGAT CGCTGAGATA GGTGCCTCAC TGATTAAGCA TTGGTAACTG TCAGACCAAG TTTACTCATA TATACTTTAG ATTGATTTAA AACTTCATTT
3001 TTAATTTAAA AGGATCTAGG TGAAGATCCT TTTTGATAAT CTCATGACCA AAATCCCTTA ACGTGAGTTT TCGTTCCACT GAGCGTCAGA CCCCGTAGAA
3101 AAGATCAAAG GATCTTCTTG AGATCCTTTT TTTCTGCGCG TAATCTGCTG CTTGCAAACA AAAAAACCAC CGCTACCAGC GGTGGTTTGT TTGCCGGATC
3201 AAGAGCTACC AACTCTTTTT CCGAAGGTAA CTGGCTTCAG CAGAGCGCAG ATACCAAATA CTGTCCTTCT AGTGTAGCCG TAGTTAGGCC ACCACTTCAA
3301 GAACTCTGTA GCACCGCCTA CATACCTCGC TCTGCTAATC CTGTTACCAG TGGCTGCTGC CAGTGGCGAT AAGTCGTGTC TTACCGGGTT GGACTCAAGA
3401 CGATAGTTAC CGGATAAGGC GCAGCGGTCG GGCTGAACGG GGGGTTCGTG CACACAGCCC AGCTTGGAGC GAACGACCTA CACCGAACTG AGATACCTAC
3501 AGCGTGAGCT ATGAGAAAGC GCCACGCTTC CCGAAGGGAG AAAGGCGGAC AGGTATCCGG TAAGCGGCAG GGTCGGAACA GGAGAGCGCA CGAGGGAGCT
3601 TCCAGGGGGA AACGCCTGGT ATCTTTATAG TCCTGTCGGG TTTCGCCACC TCTGACTTGA GCGTCGATTT TTGTGATGCT CGTCAGGGGG GCGGAGCCTA
3701 TGGAAAAACG CCAGCAACGC GGCCTTTTTA CGGTTCCTGG CCTTTTGCTG GCCTTTTGCT CACATG
```

Expression vector TAG-1-pTFT

Expression vector TAG-1-pTFT was cloned from TAG-1-pQE-60by using the restriction sites NcoI and the HindIII.

Vector map of TAG-1-pTFT:

TAG-1 (Ig1-4)	*117-1317*
bla (resistance)	*2072-2930*
ori	*3690*



DNA Sequence of human TAG-1 insert:

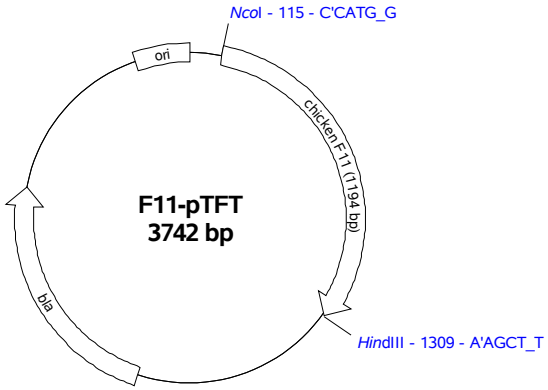
```
 101 AGAAGGAGAT ATATCCATGG CCCTGGGATC CCAAACCACC TTCGGGCCTG TCTTTGAAGA CCAGCCCCTC AGTGTGCTAT TCCCAGAGGA GTCCACGGAG
 201 GAGCAGGTGT TGCTGGCATG CCGCGCCCGG GCCAGCCCTC CAGCCACCTA TCGGTGGAAG ATGAATGGTA CCGAGATGAA GCTGGAGCCA GGTTCCCGTC
 301 ACCAGCTGGT GGGGGGCAAC CTGGTCATCA TGAACCCCAC CAAGGCACAG GATGCCGGGG TCTACCAGTG CCTGGCCTCC AACCCAGTGG GCACCGTTGT
 401 CAGCAGGGAG GCCATCCTCC GCTTCGGCTT TCTGCAGGAA TTCTCCAAGG AGGAGCGAGA CCCAGTGAAA GCTCATGAAG GCTGGGGGGT GATGTTGCCC
 501 TGTAACCCAC CTGCCCACTA CCCAGGCTTG TCCTACCGCT GGCTCCTCAA CGAGTTCCCC AACTTCATCC CGACGGACGG GCGTCACTTC GTGTCCCAGA
 601 CCACAGGGAA CCTGTACATT GCCCGAACCA ATGCCTCAGA CCTGGGCAAC TACTCCTGTT TGGCCACCAG CCACATGGAC TTCTCCACCA AGAGCGTCTT
 701 CAGCAAGTTT GCTCAGCTCA ACCTGGCTGC TGAAGATACC CGGCTCTTTG CACCCAGCAT CAAGGCCCGG TTCCCAGCAG AGACCTATGC ACTGGTGGGG
 801 CAGCAGGTCA CCCTGGAGTG CTTCGCCTTT GGGAACCCTG TCCCCCGGAT CAAGTGGCGC AAAGTGGACG GCTCCCTGTC CCCGCAGTGG ACCACAGCTG
 901 AGCCCACCCT GCAGATCCCC AGCGTCAGCT TTGAGGATGA GGGCACCTAC GAGTGTGAGG CGGAGAACTC CAAGGGCCGA GACACCGTGC AGGGCCGCAT
1001 CATCGTGCAG GCTCAGCCTG AGTGGCTAAA AGTGATCTCG GACACAGAGG CTGACATTGG CTCCAACCTG CGTTGGGGCT GTGCAGCCGC CGGCAAGCCC
1101 CGGCCTACAG TGCGCTGGCT GCGGAACGGG GAGCCTCTGG CCTCCCAGAA CCGGGTGGAG GTGTTGGCTG GGGACCTGCG GTTCTCCAAG CTGAGCCTGG
1201 AAGACTCGGG CATGTACCAG TGTGTGGCAG AGAATAAGCA CGGTACCATC TACGCCAGCG CCGAGCTAGC CGTGCAAGCA CTCGCCCCTG ACAGATCTCA
1301 TCACCATCAC CATCACTAAG CTT ...
```

Expression vector F11-pTFT

Expression vector F11-pTFT was cloned from F11-pQE-60 by using the restriction sites NcoI and the HindIII.

Vector map of F11-pTFT:

F11 (Ig1-4)	*117-1308*
bla (resistance)	*2063-2921*
ori	*3681*



DNA Sequence of F11 insert:

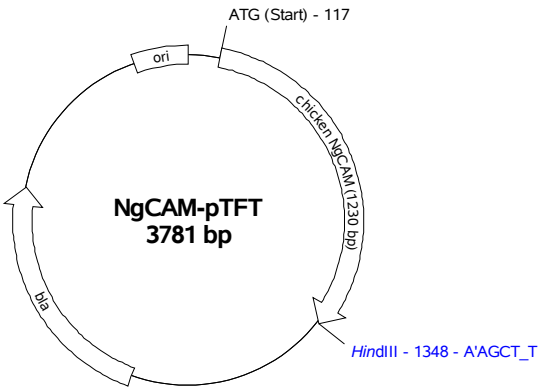
```
 101 AGAAGGAGAT ATATCCATGG CTACCCATTT TTCAGAGGAA GGAAACAAAG GTTATGGACC CGTTTTTGAG GAGCAACCCA TTGATACTAT CTATCCAGAA
 201 GAATCTTCAG ATGGACAAGT TTCAATGAAC TGCAGGGCTC GAGCAGTTCC TTTTCCTACT TACAAGTGGA AACTCAACAA CTGGGATATT GATTTAACAA
 301 AGGATCGCTA CAGTATGGTA GGAGGCAGAC TTGTTATCAG TAATCCAGAG AAATCAAGGG ATGCTGGAAA ATATGTCTGT GTAGTATCGA ATATCTTTGG
 401 AACTGTCAGA AGCAGTGAAG CCACTCTGAG TTTTGGATAT CTGGATCCTT TTCCACCAGA GGAACACTAT GAGGTTAAAG TACGAGAAGG TGTTGGAGCA
 501 GTGCTTCTTT GTGAGCCTCC TTACCACTAC CCAGATGATC TCAGTTATCG CTGGCTTCTA AATGAGTTCC CAGTATTCAT TGCTTTGGAC AGACGACGGT
 601 TTGTGTCTCA GACGAATGGC AATCTCTACA TTGCAAATGT AGAAGCATCA GATAAGGGAA ATTATTCATG TTTTGTGTCC AGCCCTTCGA TAACAAAGAG
 701 TGTATTCAGT AAATTTATTC CACTTATTCC ACAGGCCGAT CGTGCCAAAG TGTATCCTGC TGATATCAAG GTGAAGTTCA AGGACACGTA TGCTCTCCTG
 801 GGGCAAAATG TGACACTGGA GTGTTTTGCT CTTGGGAATC CGGTTCCTGA ACTCAGATGG AGTAAATACC TTGAACCTAT GCCAGCCACT GCTGAGATAA
 901 GCATGTCTGG TGCAGTTCTT AAGATCTTTA ATATTCAGTA TGAAGATGAA GGACTTTATG AGTGTGAAGC TGAAAACTAT AAAGGAAAAG ATAAACATCA
1001 AGCAAGAGTT TATGTGCAAG CCTCTCCAGA GTGGGTAGAA CATATCAACG ACACAGAGAA GGATATAGGC AGCGATCTCT ACTGGCCGTG TGTAGCTACA
1101 GGCAAGCCTA TTCCAACAAT TAGGTGGTTG AAGAACGGTG TCTCGTTCCG GAAAGGTGAA CTACGAATTC AAGGCCTGAC CTTTGAAGAT GCTGGCATGT
1201 ATCAGTGTAT AGCAGAAAAC GCACATGGAA TTATTTATGC AAATGCTGAA CTGAAGATTG TGGCTTCACC TCCTACTGTC GACAGATCTC ATCACCATCA
1301 CCATCACTAA GCTT ...
```

Expression vector NgCAM-pTFT

Expression vector NgCAM-pTFT was cloned as described (**MÖRTL, 2000**)

Vector map of NgCAM-pTFT:

NgCAM (Ig1-4)	*117-1347*
bla (resistance)	*2102-2960*
ori	*3720*



DNA Sequence of NgCAM insert:

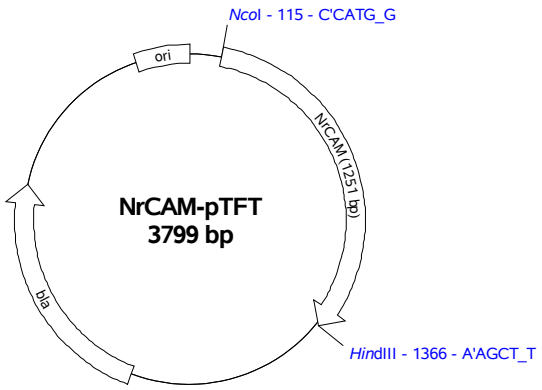
```
  101 AGAAGGAGAT ATATCCATGA TCACCATTCC CCCGGAGTAT GGTGCACACG ATTTCCTGCA GCCCCCCGAG CTGACGGAGG AACCCCCGGA ACAACTCGTG
  201 GTCTTCCCCA GTGATGACAT CGTCCTCAAA TGCGTGGCCA CCGGGAACCC CCCCGTCCAG TACCGATGGA GCCGTGAGGA TCAGCCCTTC GTCCCCGAGG
  301 AGCACGGGGG GGTCTCGGTG GTCCCCGGAT CGGGGACTTT GGTCATCAAC GCCACGTTGG CCGCGCGGCT CCAGGGGCGC TTCCGCTGCT TCGCCACCAA
  401 CGCGTTGGGC ACCGCTGTGT CTCCCGAGGC CAACGTCATC GCCGAGAACA CTCCGCAGTG GCCGGAGAAG AAGGTGACCC CGGTGGAGGT GGAGGAGGGG
  501 GACCCCGTGG TGCTGCCCTG TGACCCCCCC GAGAGCGCTG TGCCCCCTAA AATCTATTGG CTCAACAGCG ACATCGTTCA CATCGCTCAG GACGAGAGGG
  601 TATCGATGGG GCAGGATGGG AACCTCTACT TCTCCAACGC CATGGTGGGC GACAGCCACC CCGACTACAT CTGCCACGCT CACTTCCTCG GCCCCCGCAC
  701 CATCATCCAG AAGGAGCCCC TCGACCTCCG CGTGGCCCCC AGTAATGCGG TTCGGTCCCG CCGCCCCCGC CTGCTGCTGC CCCGCGACCC CCAAACGACC
  801 ACCATCGCCC TCCGGGGGGG CAGCGTCGTG TTGGAGTGCA TCGCTGAGGG GCTCCCCACT CCATGGGTCC GATGGCGGCG GCTGAACGGC CCCCTCCCCA
  901 CCCGGGCGGC GTTGGAGAAC TTCAACAAAA CGCTGCGGCT GTGGGGGGTG ACGGAGAGCG ACGACGGGGA GTACGAATGT GTGGCTGAGA ACGGGAGGGG
 1001 GACGGCCAGG GGCACCCACA GCGTCACCGT GGAGGCCGCC CCATATTGGG TGCGGCGGCC ACAGAGTGGG GTCTTCGGGC CGGGGGAGAC GGCGAGGCTG
 1101 GACTGCGAGG TGGGGGGGAA ACCCCGACCC CAAATCCAAT GGAGCATCAA TGGGGTCCCC ATCGACGCGC TGCCGGGCGC GGAGCGGCGG TGGCTGCGGG
 1201 GCGGCGCTTT GGTGCTTCCG GAGCTGCGGC CGAACGACAG CGCGGTGCTG CAGTGCGAGG CGAGGAACCG CCACGGCCGC CTATTGGCCA ACGCCTTCCT
 1301 GCACGTCGTG GAGCTGCCGG TCAGATCTCA TCACCATCAC CATCACTAAG CTT ...
```

Expression vector NrCAM-pTFT

Expression vector NrCAM-pTFT was cloned from NrCAM-pQE-60 Vector by using the restriction sites NcoI and the HindIII.

Vector map of NrCAM-pTFT:

NrCAM (Ig1-4)	*117-1366*
bla (resistance)	*2120-2978*
ori	*3738*

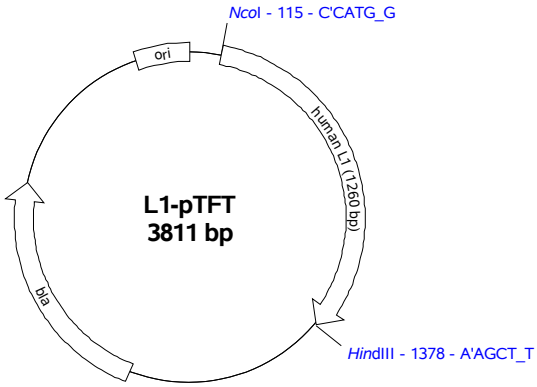


DNA Sequence of NrCAM insert:

```
  101 AGAAGGAGAT ATATCCATGG CATTGGATGT ACCTCTTGAT TCAAAACTTC TAGAAGAATT GTCTCAACCT CCAACAATAA CTCAGCAGTC TCCAAAAGAT
  201 TACATTGTTG ACCCTCGAGA GAATATTGTA ATACAATGTG AAGCAAAAGG AAAACCACCT CCTAGCTTCT CCTGGACGCG CAATGGAACT CATTTTGATA
  301 TAGATAAAGA TGCACAGGTA ACAATGAAAC CAAATTCAGG AACCCTTGTT GTAAATATTA TGAATGGTGT GAAGGCAGAA GCATATGAAG GAGTATACCA
  401 GTGTACAGCA AGGAATGAAA GAGGAGCAGC CATTTCCAAC AATATTGTTA TACGGCCATC TAGATCCCCT TTGTGGACTA AAGAAAAACT AGAACCAAAT
  501 CATGTTCGAG AAGGTGATTC CCTAGTACTA AACTGCAGAC CTCCTGTTGG CTTACCACCA CCTATAATAT TTTGGATGGA TAATGCTTTC CAAAGGCTGC
  601 CTCAAAGTGA AAGAGTTTCT CAAGGTCTCA ATGGAGACCT TTATTTTTCT AATGTACAAC CAGAGGACAC CCGTGTGGAC TATATCTGCT ACGCGAGATT
  701 TAATCACACA CAAACTATAC AGCAGAAACA ACCCATTTCT GTAAAAGTCT TTTCAACCAA GCCAGTTACA GAAAGGCCAC CAGTTCTTCT TACACCAATG
  801 GGCAGCACAA GTAACAAAGT GGAACTGAGA GGAAATGTTC TTTTGTTGGA ATGCATCGCA GCAGGATTAC CCACACCAGT AATCCGCTGG ATTAAAGAGG
  901 GTGGTGAACT GCCAGCCAAC AGAACGTTTT TTGAAAATTT TAAGAAAACT CTCAAGATTA TAGACGTCTC TGAAGCTGAC TCTGGGAACT ACAAATGTAC
 1001 AGCAAGAAAT ACATTGGGTT CTACTCATCA TGTCATTTCG GTAACTGTAA AAGCTGCCCC ATACTGGATA ACAGCACCCA GGAACTTAGT ATTGTCTCCT
 1101 GGAGAAGATG GGACATTGAT CTGCAGAGCT AATGGCAACC CAAAACCTAG CATAAGCTGG TTAACAAATG GCGTTCCCAT AGCAATTGCC CCAGAAGATC
 1201 CTAGCAGAAA GGTAGATGGG GATACCATTA TTTTCTCAGC TGTGCAAGAA CGGTCAAGTG CTGTTTATCA GTGCAATGCT TCTAATGAGT ATGGATACTT
 1301 GCTGGCAAAT GCATTTGTGA ATGTTCTTGC TGAGCCACCA AGATCTCATC ACCATCACCA TCACTAAGCT T ...
```

Expression vector L1-pTFT

Expression vector L1-pTFT was cloned from L1-pQE-60 by using the restriction sites NcoI and the HindIII.



Vector map of L1-pTFT:

L1 (Ig1-4)	*117-1377*
bla (resistance)	*2132-2990*
ori	*3750*

DNA Sequence of L1 insert:

```
 101 AGAAGGAGAT ATATCCATGG CTATCCAGAT CCCCGAGGAA TATGAAGGAC ACCATGTGAT GGAGCCACCT GTCATCACGG AACAGTCTCC ACGGCGCCTG
 201 GTTGTCTTCC CCACAGATGA CATCAGCCTC AAGTGTGAGG CCAGTGGCAA GCCCGAAGTG CAGTTCCGCT GGACGAGGGA TGGTGTCCAC TTCAAACCCA
 301 AGGAAGAGCT GGGTGTGACC GTGTACCAGT CGCCCCACTC TGGCTCCTTC ACCATCACGG GCAACAACAG CAACTTTGCT CAGAGGTTCC AGGGCATCTA
 401 CCGCTGCTTT GCCAGCAATA AGCTGGGCAC CGCCATGTCC CATGAGATCC GGCTCATGGC CGAGGGTGCC CCCAAGTGGC CAAAGGAGAC AGTGAAGCCC
 501 GTGGAGGTGG AGGAAGGGGA GTCAGTGGTT CTGCCTTGCA ACCCTCCCCC AAGTGCAGAG CCTCTCCGGA TCTACTGGAT GAACAGCAAG ATCTTGCACA
 601 TCAAGCAGGA CGAGCGGGTG ACGATGGGCC AGAACGGCAA CCTCTACTTT GCCAATGTGC TCACCTCCGA CAACCACTCA GACTACATCT GCCACGCCCA
 701 CTTCCCAGGC ATTAGGACCA TCATTCAGAA GGAACCCATT GACCTCCGGG TCAAGGCCAC CAACAGCATG ATTGACAGGA AGCCGCGCCT GCTCTTCCCC
 801 ACCAACTCCA GCACGCACCT GGTGGCCTTG CAGGGGCAGC CATTGGTCCT GGAGTGCATC GCCGAGGGCT TTCCCACGCC CACCATCAAA TCGGTGCGCC
 901 CCAGTGGCCC CATGCCAGCC GACCGTGTCA CCTACCAGAA CCACAACAAG ACCCTGCAGC TGCTGAAAGT GGGCGAGGAG GATGATGGCG AGTACCGCTG
1001 CCTGGCCGAG AACTCACTGG GCAGTGCCCG GCATGCGTAC TATGTCACCG TGGAGGCTGC CCCGTACTGG CTGCACAAGC CCCAGAGCCA TCTATATGGG
1101 CCAGGAGAGA CTGCCCGCCT GGACTGCCAA GTCGAGGGCA GGCCCCAACC AGAGGTCACC TGGAGAATCA ACGGGATCCC TGTGGAGGAG CTGGCCAAAG
1201 ACCAGAAGTA CCGGATTCAG CGTGGCGCCC TGATCCTGAG CAACGTGCAG CCCAGTGACA CAATGGTGAC CCAATGTGAG GCCCGCAACC GGCACGGGCT
1301 CTTGCTGGCC AATGCCTACA TCTACGTTGT CCAGCTGCCA GCCAAGGTCG ACAGATCTCA TCACCATCAC CATCACTAAG CTT ...
```

Expression vector E587-antigen-pTFT

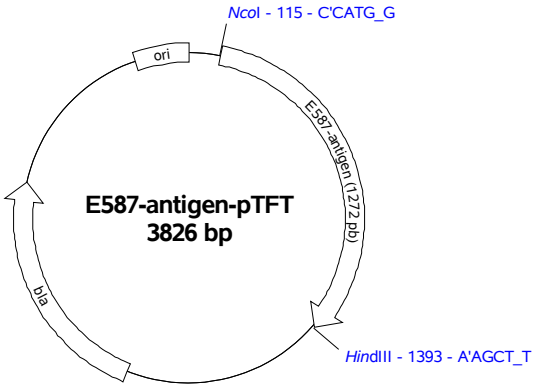
The expression vector E587-antigen-pTFT was constructed earlier by Ute Lässing of the group of Claudia Stürmer.

Vector map of E587-antigen-pTFT:

E587-antigen (Ig1-4) *117-1389*

bla (resistance) *2147-3005*

ori *3765*



DNA Sequence of E587-antigen insert:

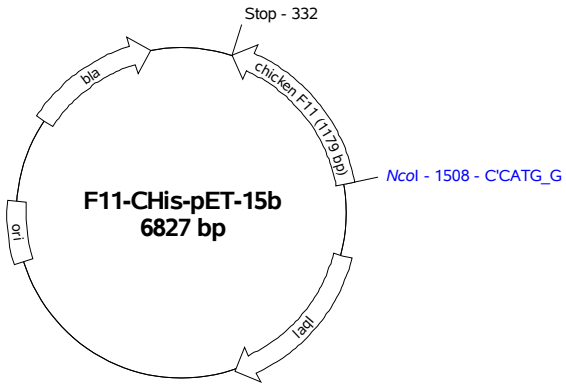
```
 101 AGAAGGAGAT ATATCCATGG CTCATCACCA CCATCACCAC ATCGAGGGAA GGATGATGCC GCCGGATTAC ACATATAGCA ACCTAAAACA AGCTCCAACG
 201 ATCACAGTTC AGCCCGTGTC TCACACTGCT TTCTCTCTGG ATGATGTGAT TCTGGCCTGT GAAGCCAGTG GAGACCCGGC ACCCAGCTTC CGCTGGGTGA
 301 AAGATGGCAA AGAGTTTAAG AGGGAACTGT TGAGTTCGGG AACTCTGACA GCAGAAGATA AAGAGGAGCT CCATCCTATC CAGGGATCGT ACCGCTGCTA
 401 CGTGATGAGC TTGGGAACCG CCGTCTCAGA CCTGGCGCAG CTGATCACAG AGCCCATCCC AACCTTGGCC AAAGAGAAGA GGCAGAAAAC AAGATCGTTT
 501 GAGGAGGGAG ACAGCGCAGT CCTCTACTGT AACCCTCCCA AGAGCTCTGT CACTCCCAAA ATACACTGGA TGGACATGCA CTGGCGCCAC ATCCCTCTGA
 601 ATGAGCGGGT GACCACTAGT CTTGATGGAA ATCTGTACTT TGCCAACTTA CTAGTCAATG ACAGCAGAGA AGACTACACC TGCAATGCCC ACATCATCAA
 701 CGCCAGCGTC ATCCTCCCTA AAGAGCGCAT CTCCATCAGC GTCACACCCT CTAATTCTGT GTTGAAGAAC CGGCGTCCCC AGCTGCAGAA GCCAGCGGGA
 801 TCCCACAGCT CTTACCTGGT GCTCAGGGGT CAAACGCTGA CTCTAGAGTG CATCCCAGAA GGCCTCCCCA CTCCAGAGGT GCAGTGGGAA CGGATGGACA
 901 GTCCTCTGTC TCCGGCGCGG GTCCGATGGC TCAAATACAA GCGCTGGCTC CAGATCGAGA GCGTGTCGGA GGCAGATGAT GGCGAGTACA CATGCACTGC
1001 CCAAAACAGC CAGGGCTCTG TTAAACACCA CTACGCTGTG ACGGTAGAGG CTGCTCCGTA CTGGACCAGG CGTCCCGAGA ACCACCTGTA CGCCCCAGGA
1101 GAGACTGTGA GGCTGGACTG TCAGGCTGAG GGGATTCCCA CCCCAAACAT CACATGGAGC ATGAACGGCG CTCCCATCGC AGGGACAGAC CCGGACCCGA
1201 GACGGCATGT GAGCTCAGGG ACACTGATCC TGACTGACGT TCAGATCAGC GATACGGCGG TCTATCATGT CGAGGCCACC AATAAACACG GCAATATCCT
1301 GATCAACACA CATGTCCATG TAGTTGAGCT GCCTCCTCAG ATCCTGACAG AAGACGATCT CAAGTACGAG GCCACAGAAG GGCAAACGTA GTAAGCTT ..
```

Expression vector F11-CHis-pET-15b

Vektor F11-CHis-pET-15b was cloned from cDNA with primers F11-N-nativ (*cacacccatggc-tacccatttttcagaggaagg*; NcoI), F11-C-6HIS (*gtgtggtcgacttagtggtggtggtggtggtgagtaggaggtgaa-gccac*; SalI), and vector pET-15b (NcoI and XhoI).

Vector map of F11-CHis-pET-15b:

F11 (Ig1-4)	*1511-332*
bla (resistance)	*5762-6619*
laqI	*1985-3064*
ori	*5001*



DNA Sequence of F11-CHis-pET-15b insert :

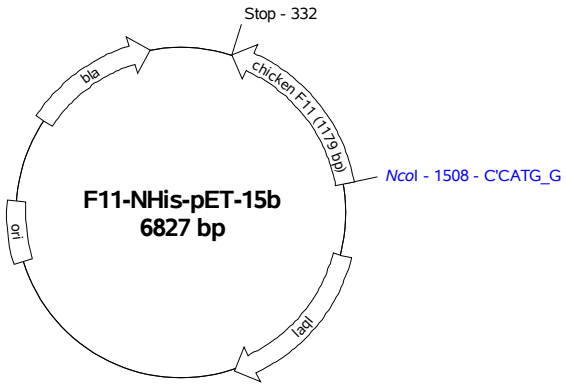
```
  311 TAGCAGCCGG ATCCTCGACT TAGTGGTGGT ...
  ...(F11 Ig1-4)...
 1491 CTCTGAAAAA TGGGTAGCCA TGGTATATCT
```

Expression vector F11-NHis-pET-15b

Vektor F11-NHis-pET-15b was cloned from cDNA with primers F11-N-6HIS (*cacacccatggc-tcaccaccaccaccaccacacccatttttcagaggaagg*; NcoI), F11-C-nativ (*gtgtggtcgacttaagtaggaggt-gaagccac*; SalI), and vector pET-15b (NcoI and XhoI).

Vector map of F11-NHis-pET-15b:

F11 (Ig1-4)	*1511-332*
bla (resistance)	*5762-6619*
laqI	*1985-3064*
ori	*5001*



DNA Sequence of F11-NHis-pET-15b insert :

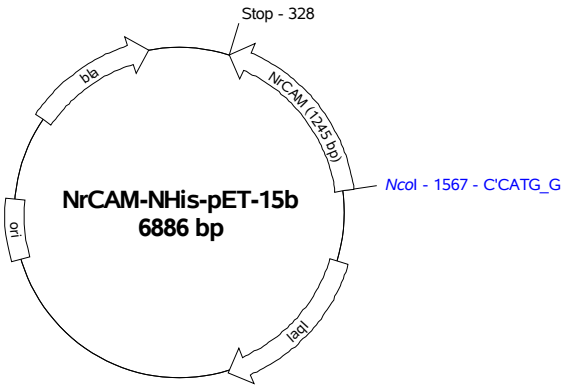
```
 311 TAGCAGCCGG ATCCTCGACT TAAGTAGGAG  ...
 ...(F11 Ig1-4)...
1491 GTGGTGGTGG TGGTGAGCCA TGGTATATCT
```

Expression vector NrCAM-NHis-pET-15b

Vector NrCAM-NHis-pET-15b was cloned from cDNA with primers NrCAM-N-6HIS (*cacacccatggcacaccaccaccaccaccacttggatgtacctcttgattc*; NcoI), NrCAM-C-nativ (*gtgtgagatctttatggtggctcagcaagaacattc*; BglII), and vector pET-15b (NcoI and BamHI).

Vector map of NrCAM-NHis-pET-15b:

NrCAM (Ig1-4)	*1570-328*
bla (resistance)	*5821-6678*
laqI	*2044-3123*
ori	*5060*



DNA Sequence of NrCAM-NHis-pET-15b insert :

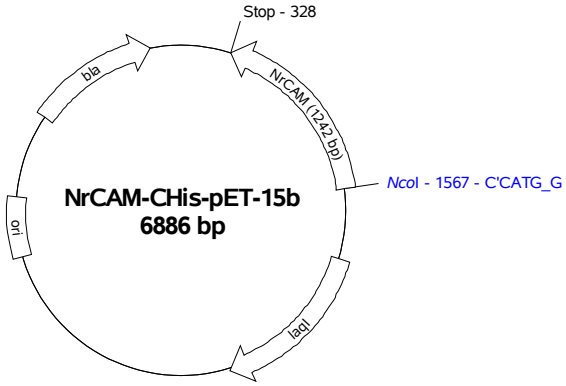
```
  311 TAGCAGCCGG ATCTTTTATG GTGGCTCAGC
  ...(NrCAM Ig1-4)...
 1551 TGGTGGTGGT GGTGTGCCAT GGTATATCTC
```

Expression vector NrCAM-CHis-pET-15b

Vector NrCAM-CHis-pET-15b was cloned from cDNA with primers NrCAM-C-6HIS (gtg-tgctcgagttagtggtggtggtggtggtgagatctgtcaggggcgagtgcttgc; BglII (a'gatct), NrCAM-N-nativ (*cacacccatggccctgggatcccaaaccac*; NcoI), and vector pET-15b (NcoI and BamHI (g'gatcc)).

Vector map of NrCAM-CHis-pET-15b:

NrCAM (Ig1-4)	*1570-328*
bla (resistance)	*5821-6678*
laqI	*2044-3123*
ori	*5060*



DNA Sequence of NrCAM-CHis-pET-15b insert :

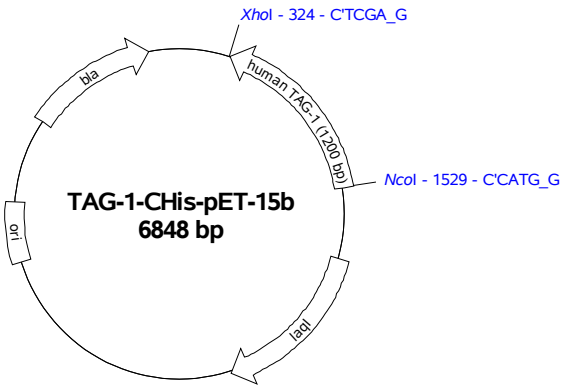
```
  311 TAGCAGCCGG ATCTTTTAGT GGTGGTGGTG ...
  ...(NrCAM Ig1-4)...
 1551 AGAGGTACAT CCAATGCCAT GGTATATCTC
```

Expression vector TAG-1-CHis-pET-15b

Vector TAG-1-CHis-pET-15b was cloned from cDNA with primers TAG-1-CHIS (*gtgtgctcg-agttagtggtggtggtggtgagatctgtcaggggcgagtgcttgc*; NcoI) and TAG-1-N-nativ (*cacacccat-ggccctgggatcccaaaccac*, XhoI), and vector pET-15b (NcoI, XhoI).

Vector map of TAG-1-CHis-pET-15b:

TAG-1 (Ig1-4)	*1532-332*
bla (resistance)	*5783-6640*
laqI	*2006-3085*
ori	*5022*



DNA Sequence of TAG-1 insert :

```
311 TAGCAGCCGG ATCCTCGAGT TAGTGGTGGT GGTGGTGGTG ...
...(TAG-1 Ig1-4)...
1511 TGGTTTGGGA TCCCAGGGCC ATGGTATATC
```

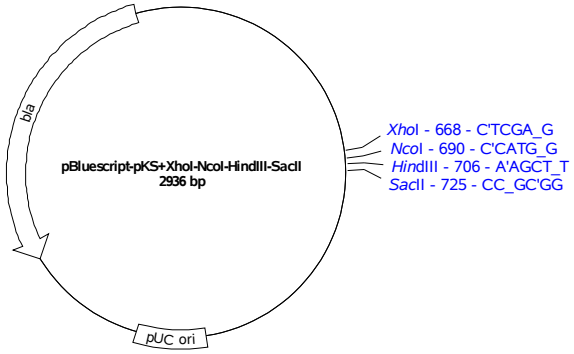
Cloning vector pBluscript II SK+ modified for pPICZαA generation

The vector was obtained by ligation of to annealed primers, to obtain a modified multiple cloning site. The vector was treated with the restriction enzymes XhoI and SacII. Primers are given below.

Vector map of modified pBluescript II SK+

bla (resistance) *1948-2806*

ori *1500*



Detail of multiple cloning site:

```
   XhoI                    NcoI            HindIII              SacII
   |                       |               |                    |
CCC'TCGA_GAAAAGAGAGGCTGAAGCCATGGCACACTGTGTAAGCTTAATTAGCTGACC_GC'GGTG
GGG_AGCT'CTTTTCTCTCCGACTTCGGTACCGTGTGACACATTCGAATTAATCGACTGG'CG_CCAC
```

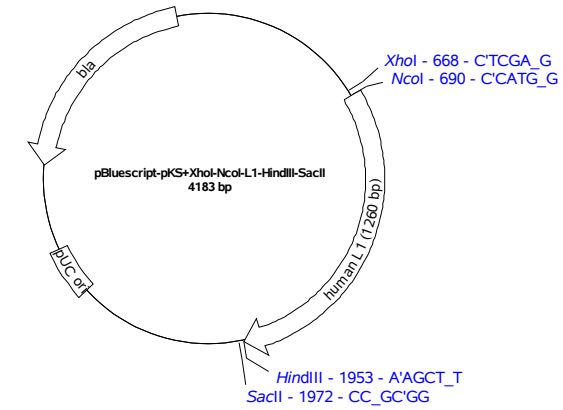
Forward primer: 5'-*tcgagaaaagagaggctgaagccatggcacactgtgtaagcttaattagctgaccgc-3'*
Backward primer: 5'-*ccagtcgattaattcgaatgtgtcacacggtaccgaagtcggagagaaaag-3'*

Expression vector L1-pPICZαA

For the construction of L1-pPICZαA a modified pBluscript vector was generated, by ligation of an L1-insert obtained by NcoI/HindIII digestion of L1-pQE-60 into the modified pBluscript-II-SK+ vector.

Vector map of L1-pBluescript-II-KS+:

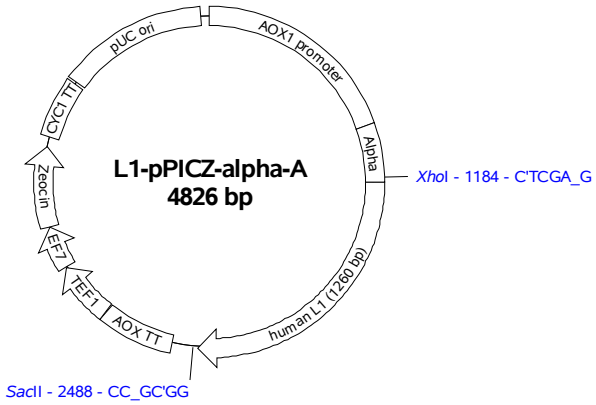
L1 (Ig1-4)	*692- 1952*
bla (resistance)	*3195-4053*
ori	*2747*



The L1-pPICZαAvector was then obtained by ligation of the L1-insert from the pBluscript vector, using the restriction sites XhoI and SacII.

Vector map of L1-pPICZαA:

AOX1 promoter	*1-946*
Alpha signal sequence	*941-1208*
L1 (Ig1-4)	*1208-2468*
AOX1 termination	*2588-2929*
TEF1 promoter	*2930-3207*
EF1 promoter	*3208-3408*
Zeocin (resistance)	*3410-3784*
CYC1 termination	*3830-4102*
pUC origin	*4113-4786*

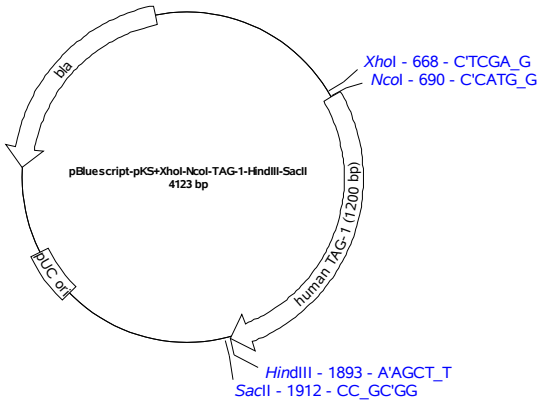


Expression vector TAG-1-pPICZαA

For the construction of TAG1-pPICZαA a modified pBluscript vector was generated, by ligation of an TAG-1-insert obtained by NcoI/HindIII digestion of TAG-1-pQE-60 into the modified pBluscript-II-SK+ vector.

Vector map of TAG-1-pBluescript-II-KS+:

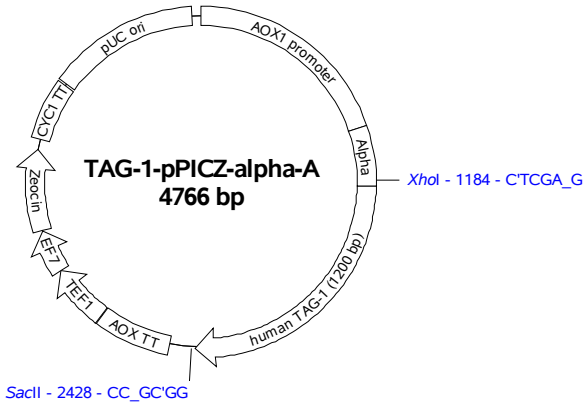
TAG-1 (Ig1-4)	*692- 1892*
bla (resistance)	*3165-3993*
ori	*2687*



The TAG-1-pPICZαAvector was then obtained by ligation of the TAG-1-insert from the pBluscript vector, using the restriction sites XhoI and SacII.

Vector map of TAG-1-pPICZαA:

AOX1 promoter	*1-946*
Alpha signal sequence	*941-1208*
TAG-1 (Ig1-4)	*1208-2408*
AOX1 termination	*2528-2869*
TEF1 promoter	*2870-3147*
EF1 promoter	*3148-3348*
Zeocin (resistance)	*3350-3724*
CYC1 termination	*3770-4042*
pUC origin	*4053-4726*

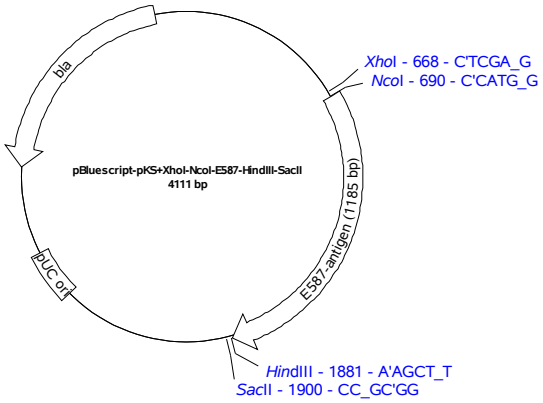


Expression vector E587-pPICZαA

For the construction of E587-pPICZαA a modified pBluescript vector was generated, by ligation of an E587-insert obtained by PCR with the primers E587-C-PIC (*gtgtgaagcttactactc-aactacatggacatgtgtg*; HindIII) and E587-NHis-PIC (*cacacccatggctcatcaccaccatcaccacatcga-gggaaggctaaaacaagctccaacgatc*; NcoI), using the E587-antigen-pTFT expression vector as template for PCR. The E587-antigen fragment is truncated compared to the E587-antigen-pTFT: 10 residues are deleted from the N-terminal part and 20 amino acids from the C-terminal part. The corresponding amino acid sequence is given in an multiple alignment (see page 60)

Map of E587-antigen-pBluescript-II-KS+:

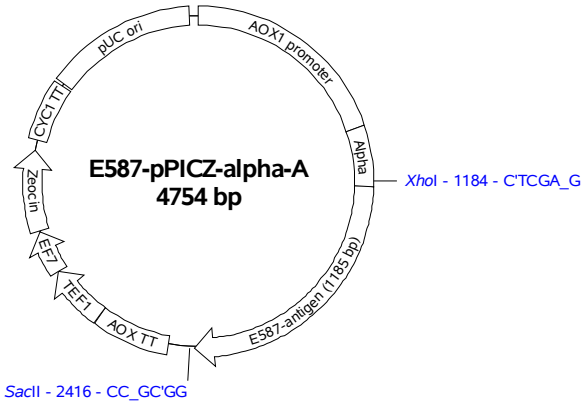
E587-antigen (Ig1-4)	*692-1877*
bla (resistance)	*3123*
ori	*2675*



The e587-antigen-pPICZαAvector was then obtained by ligation of the E587-antigen-insert from the pBluscript vector, using the restriction sites XhoI and SacII.

Vector map of E587-antigen-pPICZαA:

AOX1 promoter	*1-946*
Alpha signal sequence	*941-1208*
E587-antigen (Ig1-4)	*1208-2393*
AOX1 termination	*2516-2857*
TEF1 promoter	*2858-3135*
EF1 promoter	*3136-3336*
Zeocin (resistance)	*3338-3712*
CYC1 termination	*3758-4030*
pUC origin	*4041-4714*

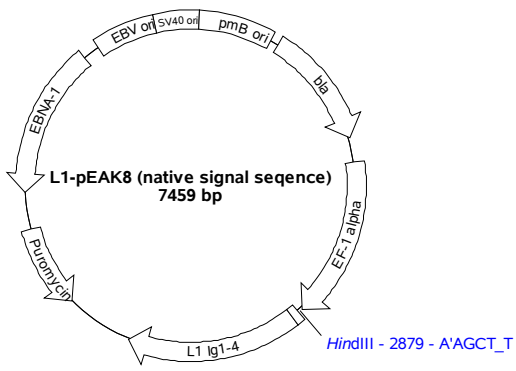


HEK cell expression vector L1-pEAK8 (native signal sequence)

The expression vector was cloned from cDNA with primers L1-N-HindIII-pEAK8 (*cacaca-agcttccagccatggtagtggtgctgcggtacg*; HindIII) and L1-CHis-pEAK8 (*cacactcaatgatgatgatga-tgatgcttggctggcagctggaca*; blunt end) into the pEAK8 vektor (Edge BioSystems, USA) using the restriction sites HindIII and EcoRV.

Vector map of L1-pEAK8 (native signal sequence)

pmB origin	*60-699*
bla (resistance)	*659-1517*
EF-1 alpha	*1695-2860*
L1 signal sequence	*2891-2948*
L1 (Ig1-4)	*2948-4190*
Puromycin (resistance)	*5280-4683*
EBNA-1	*6636-5544*
EBV origin	*6757-7195*
SV40 origin	*7195-60*



Vector sequence at the L1 signal sequence:

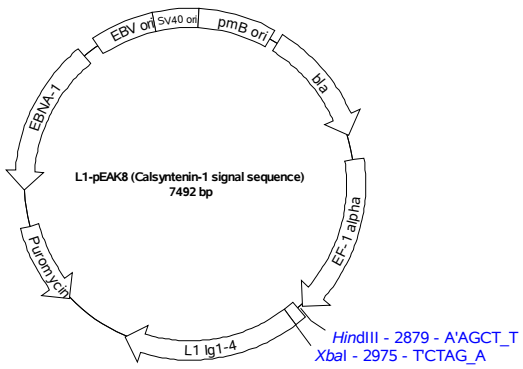
```
                                  HindIII
     .         .         .       | .         .         .         .
2851 ttttttcttccatttcaggtgtcgtgaaaagcttccagccatggtagtggtgctgcggtac
                                                M  V  V  V  L  R  Y

                                                   L1 Ig-1 Start
              .         .         .         .    |     .
2902  gtgtggcctctcctcctctgcagcccctgcctgcttatccagatccccgaggaat ...
       V  W  P  L  L  L  C  S  P  C  L  L  I  Q  I  P  E  E   ...
```

HEK cell expression vector L1-pEAK8 (Calsyntenin-1 signal sequence)

The expression vector was cloned from L1 cDNA with primers L1-N-XbaI-pEAK8 (*cacact-tagaatccagatccccgaggaatatg*; XbaI) and L1-CHis-pEAK8 (*cacactcaatgatgatgatgatgcttg-gctggcagctggaca*; blunt end) into the pEAK8-Calsyntenin-1 vector, provided by Kerstin Leuthäuser from the group of P. Sonderegger (University of Zürich). The XbaI and EcoRV restriction sites were used. The XbaI restriction sites is placed at the end of the Calsyntenin-1 signal sequence, thus enabling a fusion of the L1 fragment to the Calsyntenin-1 signal sequence.

Vector map of L1-pEAK8 (Calsyntenin-1 signal sequence)

pmB origin	*60-599*
bla (resistance)	*659-1517*
EF-1 alpha	*1695-2860*
Calsyntenin-1 signal seq.	*2891-2975*
L1 (Ig1-4)	*2976-4229*
Puromycin (resistance)	*5313-4716*
EBNA-1	*6669-5577*
EBV origin	*6790-7228*
SV40 origin	*7228-60*



Vector sequence at the Calsyntenin-1 signal sequence:

```
                                  HindIII
     .         .         .        | .         .         .         .
2851 ttttttcttccatttcaggtgtcgtgaaaagcttccagccatgctgcgccgccccgctccc
                                              M   L   R   R   P   A   P


               .         .         .         .         .         .
2902  gcgctggccccggccgcccggctgctgctggccgggctgctgtgcggcggcggggtctgg
       A   L   A   P   A   A   R   L   L   L   A   G   L   L   C   G   G   G   V   W

                 L1 Ig-1 Start
                .|         .
2952  gcctctagaatccagatccccgaggaata ...
       A   S   R   I   Q   I   P   E   E     ...
```

References

ABEL, E. L. & SOKOL, R.J. (1991) A revised conservative estimate of the incidence of FAS and its economic impact. *Alcohol Clin Exp Res*, **15**, 514-24.

ALVAREZ-DOLADO, M., FIGUEROA, A., KOZLOV, S., SONDEREGGER, P., FURLEY, A. J. AND MUÑOZ, A. (2001) Thyroid hormone regulates TAG-1 expression in the developing rat brain. *Eur J Neurosci*, **14**, 1209-18.

BACHMANN, B. J. (1972) Pedigrees of some mutant strains of Escherichia coli K-12. *Bacteriol Rev*, **36**, 525-57.

BAHADUR, R. P., CHAKRABARTI, P., RODIER, F. AND JANIN, J. (2004) A dissection of specific and non-specific protein-protein interfaces. *J Mol Biol*, **336**, 943-55.

BANEYX, F. (1999) Recombinant protein expression in Escherichia coli. *Curr Opin Biotechnol*, **10**, 411-21.

BASS, S., GU, Q. & CHRISTEN, A. (1996) Multicopy suppressors of prc mutant Escherichia coli include two HtrA (DegP) protease homologs (HhoAB), DksA, and a truncated R1pA. *J Bacteriol*, **178**, 1154-61.

BAZAN, J. F. (1990) Structural design and molecular evolution of a cytokine receptor superfamily. *Proc Natl Acad Sci U S A*, **87**, 6934-8.

BERMAN, H. M., WESTBROOK, J., FENG, Z., GILLILAND, G., BHAT, T. N., WEISSIG, H., SHINDYALOV, I. N. AND BOURNE, P.E. (2000) The Protein Data Bank. *Nucleic Acids Res*, **28**, 235-42.

BIEBER, A. J., SNOW, P. M., HORTSCH, M., PATEL, N. H., JACOBS, J. R., TRAQUINA, Z. R., SCHILLING, J. & GOODMAN, C.S. (1989) Drosophila neuroglian: a member of the immunoglobulin superfamily with extensive homology to the vertebrate neural adhesion molecule L1. *Cell*, **59**, 447-60.

BORK, P., HOLM, L. & SANDER, C. (1994) The immunoglobulin fold. Structural classification, sequence patterns and common core. *J Mol Biol*, **242**, 309-20.

BRADFORD, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem*, **72**, 248-54.

BRÜMMENDORF, T. AND RATHJEN, F.G. (1996) Structure/function relationships of axon-associated adhesion receptors of the immunoglobulin superfamily. *Curr Opin Neurobiol*, **6**, 584-93.

BRÜMMENDORF, T., KENWRICK, S. & RATHJEN, F.G. (1998) Neural cell recognition molecule L1: from cell biology to human hereditary brain malformations. *Curr Opin Neurobiol*, **8**, 87-97.

BRÜMMENDORF, T., WOLFF, J. M., FRANK, R. AND RATHJEN, F. (1989) Neural cell recognition molecule F11: homology with fibronectin type III and immunoglobulin type C domains. *Neuron*, **2**, 1351-61.

BRÜNGER, A. T., ADAMS, P. D., CLORE, G. M., DELANO, W. L., GROS, P., GROSSE-KUNSTLEVE, R.

W., JIANG, J. S., KUSZEWSKI, J., NILGES, M., PANNU, N. S., READ, R. J., RICE, L. M., SIMONSON, T. & WARREN, G. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr*, **54 (Pt 5)**, 905-21.

BUCHNER, J. & RUDOLPH, R. (1991) Renaturation, purification and characterization of recombinant Fab-fragments produced in Escherichia coli. *Biotechnology (N Y)*, **9**, 157-62.

BUCHSTALLER, A., KUNZ, S., BERGER, P., KUNZ, B., ZIEGLER, U., RADER, C. AND SONDEREGGER, P. (1996) Cell adhesion molecules NgCAM and axonin-1 form heterodimers in the neuronal membrane and cooperate in neurite outgrowth promotion. *J Cell Biol*, **135**, 1593-607.

BURGOON, M. P., GRUMET, M., MAURO, V., EDELMAN, G. M. & CUNNINGHAM, B.A. (1991) Structure of the chicken neuron-glia cell adhesion molecule, Ng-CAM: origin of the polypeptides and relation to the Ig superfamily. *J Cell Biol*, **112**, 1017-29.

BURGOON, M. P., GRUMET, M., MAURO, V., EDELMAN, G. M. AND CUNNINGHAM, B.A. (1991) Structure of the chicken neuron-glia cell adhesion molecule, Ng-CAM: origin of the polypeptides and relation to the Ig superfamily. *J Cell Biol*, **112**, 1017-29.

BURLEY, S. K. AND PETSKO, G.A. (1985) Aromatic-aromatic interaction: a mechanism of protein structure stabilization. *Science*, **229**, 23-8.

BUTTIGLIONE, M., REVEST, J. M., PAVLOU, O., KARAGOGEOS, D., FURLEY, A., ROUGON, G. AND FAIVRE-SARRAILH, C. (1998) A functional interaction between the neuronal adhesion molecules TAG-1 and F3 modulates neurite outgrowth and fasciculation of cerebellar granule cells. *J Neurosci*, **18**, 6853-70.

CHOTHIA, C. AND JONES, E.Y. (1997) The molecular structure of cell adhesion molecules. *Annu Rev Biochem*, **66**, 823-62.

CLARKE, J., COTA, E., FOWLER, S. B. AND HAMILL, S.J. (1999) Folding studies of immunoglobulin-like beta-sandwich proteins suggest that they share a common folding pathway. *Structure*, **7**, 1145-53.

COLLABORATIVE COMPUTATIONAL PROJECT NUMBER 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr*, **50**, 760-3.

DE ANGELIS, E., MACFARLANE, J., DU, J. S., YEO, G., HICKS, R., RATHJEN, F. G., KENWRICK, S. & BRÜMMENDORF, T. (1999) Pathological missense mutations of neural cell adhesion molecule L1 affect homophilic and heterophilic binding activities. *EMBO J*, **18**, 4744-53.

DE ANGELIS, E., MACFARLANE, J., DU, J. S., YEO, G., HICKS, R., RATHJEN, F. G., KENWRICK, S. AND BRÜMMENDORF, T. (1999) Pathological missense mutations of neural cell adhesion molecule L1 affect homophilic and heterophilic binding activities. *EMBO J*, **18**, 4744-53.

DE ANGELIS, E., WATKINS, A., SCHAFER, M., BRUMMENDORF, T. & KENWRICK, S. (2002) Disease-associated mutations in L1 CAM interfere with ligand interactions and cell-surface expression. *Hum Mol Genet*, **11**, 1-12.

DENAXA, M., CHAN, C. H., SCHACHNER, M., PARNAVELAS, J. G. AND KARAGOGEOS, D. (2001) The adhesion molecule TAG-1 mediates the migration of cortical interneurons from the ganglionic eminence along the corticofugal fiber system. *Development*, **128**, 4635-44.

DEO, Y. M., GRAZIANO, R. F., REPP, R. & VAN DE WINKEL, J.G. (1997) Clinical significance of IgG Fc receptors and Fc gamma R-directed immunotherapies. *Immunol Today*, **18**, 127-35.

DEREWENDA, Z. S. (2004) Rational protein crystallization by mutational surface engineering. *Structure*, **12**, 529-35.

DIEDERICHS, K. (1995) Structural superposition of proteins with unknown alignment and detection of topological similarity using a six-dimensional search algorithm. *Proteins*, **23**, 187-95.

DIEDERICHS, K. & KARPLUS, P.A. (1997) Improved R-factors for diffraction data analysis in macromolecular crystallography. *Nat Struct Biol*, **4**, 269-75.

DIEDERICHS, K. AND KARPLUS, P.A. (1997) Improved R-factors for diffraction data analysis in macromolecular crystallography. *Nat Struct Biol*, **4**, 269-75.

DODD, J. & JESSELL, T.M. (1988a) Axon guidance and the patterning of neuronal projections in vertebrates. *Science*, **242**, 692-9.

DODD, J. AND JESSELL, T.M. (1988b) Axon guidance and the patterning of neuronal projections in vertebrates. *Science*, **242**, 692-9.

DODD, J., MORTON, S. B., KARAGOGEOS, D., YAMAMOTO, M. & JESSELL, T.M. (1988) Spatial regulation of axonal glycoprotein expression on subsets of embryonic spinal neurons. *Neuron*, **1**, 105-16.

DODD, J., MORTON, S. B., KARAGOGEOS, D., YAMAMOTO, M. AND JESSELL, T.M. (1988) Spatial regulation of axonal glycoprotein expression on subsets of embryonic spinal neurons. *Neuron*, **1**, 105-16.

DYM, O. & EISENBERG, D. (2001) Sequence-structure analysis of FAD-containing proteins. *Protein Sci*, **10**, 1712-28.

EMSLEY, P. AND COWTAN, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*, **60**, 2126-32.

ENTSCH, B. & VAN BERKEL, W.J. (1995) Structure and mechanism of para-hydroxybenzoate hydroxylase. *FASEB J*, **9**, 476-83.

FALK, J., BONNON, C., GIRAULT, J. & FAIVRE-SARRAILH, C. (2002) F3/contactin, a neuronal cell adhesion molecule implicated in axogenesis and myelination. *Biol Cell*, **94**, 327-34.

FELSENFELD, D. P., HYNES, M. A., SKOLER, K. M., FURLEY, A. J. AND JESSELL, T.M. (1994) TAG-1 can mediate homophilic binding, but neurite outgrowth on TAG-1 requires an L1-like molecule and beta 1 integrins. *Neuron*, **12**, 675-90.

FISHER, S. (1993). Utilization of amino acids and other nitrogen-containing compounds. A. SONENSHEIN, J. HOCH & R. LOSICK (EDS.). *Bacillus subtilis and Other Gram-positive Bacteria: Biochemistry, Physiology, and Molecular Biology*. pp. 221-228. , Washington, D. C.

FISHER, S. H. (1999) Regulation of nitrogen metabolism in Bacillus subtilis: vive la difference!. *Mol Microbiol*, **32**, 223-32.

FITZLI, D., STOECKLI, E. T., KUNZ, S., SIRIBOUR, K., RADER, C., KUNZ, B., KOZLOV, S. V.,

BUCHSTALLER, A., LANE, R. P., SUTER, D. M., DREYER, W. J. & SONDEREGGER, P. (2000) A direct interaction of axonin-1 with NgCAM-related cell adhesion molecule (NrCAM) results in guidance, but not growth of commissural axons. *J Cell Biol*, **149**, 951-68.

FITZLI, D., STOECKLI, E. T., KUNZ, S., SIRIBOUR, K., RADER, C., KUNZ, B., KOZLOV, S. V., BUCHSTALLER, A., LANE, R. P., SUTER, D. M., DREYER, W. J. AND SONDEREGGER, P. (2000) A direct interaction of axonin-1 with NgCAM-related cell adhesion molecule (NrCAM) results in guidance, but not growth of commissural axons. *J Cell Biol*, **149**, 951-68.

FRANSEN, E., VAN CAMP, G., VITS, L. & WILLEMS, P.J. (1997) L1-associated diseases: clinical geneticists divide, molecular geneticists unite. *Hum Mol Genet*, **6**, 1625-32.

FREIGANG, J. (1999) Strukturanalyse eines aminoterminalen Fragments des neuronalen Zelladhäsionsproteins Axonin-1 aus dem Huhn. *Universität Konstanz*, Dissertation.

FREIGANG, J., PROBA, K., LEDER, L., DIEDERICHS, K., SONDEREGGER, P. & WELTE, W. (2000) The crystal structure of the ligand binding module of axonin-1/TAG-1 suggests a zipper mechanism for neural cell adhesion. *Cell*, **101**, 425-33.

FREIGANG, J., PROBA, K., LEDER, L., DIEDERICHS, K., SONDEREGGER, P. AND WELTE, W. (2000) The crystal structure of the ligand binding module of axonin-1/TAG-1 suggests a zipper mechanism for neural cell adhesion. *Cell*, **101**, 425-33.

GENNARINI, G., CIBELLI, G., ROUGON, G., MATTEI, M. G. AND GORIDIS, C. (1989) The mouse neuronal cell surface protein F3: a phosphatidylinositol-anchored member of the immunoglobulin superfamily related to chicken contactin. *J Cell Biol*, **109**, 775-88.

GILL, S. C. & VON HIPPEL, P.H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem*, **182**, 319-26.

GILL, S. C. AND VON HIPPEL, P.H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem*, **182**, 319-26.

GIORDANO, S., LAESSING, U., ANKERHOLD, R., LOTTSPEICH, F. & STUERMER, C.A. (1997) Molecular characterization of E587 antigen: an axonal recognition molecule expressed in the goldfish central nervous system. *J Comp Neurol*, **377**, 286-97.

GIORDANO, S., LAESSING, U., ANKERHOLD, R., LOTTSPEICH, F. AND STUERMER, C.A. (1997) Molecular characterization of E587 antigen: an axonal recognition molecule expressed in the goldfish central nervous system. *J Comp Neurol*, **377**, 286-97.

GRUMET, M. AND EDELMAN, G.M. (1984) Heterotypic binding between neuronal membrane vesicles and glial cells is mediated by a specific cell adhesion molecule. *J Cell Biol*, **98**, 1746-56.

GRUMET, M., MAURO, V., BURGOON, M. P., EDELMAN, G. M. AND CUNNINGHAM, B.A. (1991) Structure of a new nervous system glycoprotein, Nr-CAM, and its relationship to subgroups of neural cell adhesion molecules. *J Cell Biol*, **113**, 1399-412.

GUPTA, R., JUNG, E. & BRUNAK, S. (2004) Prediction of N-glycosylation sites in human proteins, ***In preparation***.

GUSTAFSSON, C., GOVINDARAJAN, S. AND MINSHULL, J. (2004) Codon bias and heterologous protein expression. *Trends Biotechnol*, **22**, 346-53.

HABER, E. & ANFINSEN, C.B. (1961) Regeneration of enzyme activity by air oxidation of reduced subtilisin-modified ribonuclease. *J Biol Chem*, **236**, 422-4.

HASLER, T. H., RADER, C., STOECKLI, E. T., ZUELLIG, R. A. AND SONDEREGGER, P. (1993) cDNA cloning, structural features, and eucaryotic expression of human TAG-1/axonin-1. *Eur J Biochem*, **211**, 329-39.

HASPEL, J., SCHURMANN, G., JACOB, J., ERICKSON, H. P. AND GRUMET, M. (2001) Disulfide-mediated dimerization of L1 Ig domains. *J Neurosci Res*, **66**, 347-55.

HAYWARD, S. AND LEE, R.A. (2002) Improvements in the analysis of domain motions in proteins from conformational change: DynDom version 1.50. *J Mol Graph Model*, **21**, 181-3.

HU, Q., ANG, B., KARSAK, M., HU, W., CUI, X., DUKA, T., TAKEDA, Y., CHIA, W., SANKAR, N., NG, Y., LING, E., MACIAG, T., SMALL, D., TRIFONOVA, R., KOPAN, R., OKANO, H., NAKAFUKU, M., CHIBA, S., HIRAI, H., ASTER, J. C., SCHACHNER, M., PALLEN, C. J., WATANABE, K. & XIAO, Z. (2003) F3/contactin acts as a functional ligand for Notch during oligodendrocyte maturation. *Cell*, **115**, 163-75.

HU, Q., MA, Q., GENNARINI, G. & XIAO, Z. (2006) Cross-talk between F3/contactin and Notch at axoglial interface: a role in oligodendrocyte development. *Dev Neurosci*, **28**, 25-33.

HUECK, C. J. & HILLEN, W. (1995) Catabolite repression in Bacillus subtilis: a global regulatory mechanism for the gram-positive bacteria?. *Mol Microbiol*, **15**, 395-401.

HYNES, R. O. & LANDER, A.D. (1992) Contact and adhesive specificities in the associations, migrations, and targeting of cells and axons. *Cell*, **68**, 303-22.

HYNES, R. O. AND LANDER, A.D. (1992) Contact and adhesive specificities in the associations, migrations, and targeting of cells and axons. *Cell*, **68**, 303-22.

INVITROGEN (2004) One Shot TOP10 Competent Cells. **Version M**, *Invitrogen Corporation*, USA

INVITROGEN (2005) EasySelect Pichia Expression Kit - A Manual of Methods for Expression of Reconbinant Proteins Using pPICZ and pPICZalpha in Pichia Pastoris. **Version H**, *Invitrogen Corporation*, USA

INVITROGEN (2005) EasySelect Pichia Expression Kit - A Manual of Methods for Expression of Reconbinant Proteins Using pPICZ and pPICZalpha in Pichia Pastoris. **Version H**, *Invitrogen Corporation*, USA

ITOH, K., FUSHIKI, S., KAMIGUCHI, H., ARNOLD, B., ALTEVOGT, P. & LEMMON, V. (2005) Disrupted Schwann cell-axon interactions in peripheral nerves of mice with altered L1-integrin interactions. *Mol Cell Neurosci*, **30**, 131-6.

JANIN, J. (1997) Specific versus non-specific contacts in protein crystals. *Nat Struct Biol*, **4**, 973-4.

JANIN, J. AND RODIER, F. (1995) Protein-protein interaction at crystal contacts. *Proteins*, **23**, 580-7.

JAROSZEWSKI, L., RYCHLEWSKI, L., LI, Z., LI, W. & GODZIK, A. (2005) FFAS03: a server for

profile--profile sequence alignments. *Nucleic Acids Res*, **33**, W284-8.

JESSELL, T. M. (1988) Adhesion molecules and the hierarchy of neural development. *Neuron*, **1**, 3-13.

JOB, V., MARCONE, G. L., PILONE, M. S. & POLLEGIONI, L. (2002a) Glycine oxidase from Bacillus subtilis. Characterization of a new flavoprotein. *J Biol Chem*, **277**, 6985-93.

JOB, V., MOLLA, G., PILONE, M. S. & POLLEGIONI, L. (2002b) Overexpression of a recombinant wild-type and His-tagged Bacillus subtilis glycine oxidase in Escherichia coli. *Eur J Biochem*, **269**, 1456-63.

JONES, E. Y., DAVIS, S. J., WILLIAMS, A. F., HARLOS, K. & STUART, D.I. (1992) Crystal structure at 2.8 A resolution of a soluble form of the cell adhesion molecule CD2. *Nature*, **360**, 232-9.

KABSCH, W. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J Appl Cryst*, **26**, 795-800.

KABSCH, W. & SANDER, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577-637.

KAMEI, Y., TSUTSUMI, O., TAKETANI, Y. AND WATANABE, K. (1998) cDNA cloning and chromosomal localization of neural adhesion molecule NB-3 in human. *J Neurosci Res*, **51**, 275-83.

KARAGOGEOS, D., MORTON, S. B., CASANO, F., DODD, J. AND JESSELL, T.M. (1991) Developmental expression of the axonal glycoprotein TAG-1: differential regulation by central and peripheral neurons in vitro. *Development*, **112**, 51-67.

KARAGOGEOS, D., POURQUIÉ, C., KYRIAKOPOULOU, K., TAVIAN, M., STALLCUP, W., PÉAULT, B. AND POURQUIÉ, O. (1997) Expression of the cell adhesion proteins BEN/SC1/DM-GRASP and TAG-1 defines early steps of axonogenesis in the human spinal cord. *J Comp Neurol*, **379**, 415-27.

KARPLUS, P. A. & SCHULZ, G.E. (1987) Refined structure of glutathione reductase at 1.54 A resolution. *J Mol Biol*, **195**, 701-29.

KENWRICK, S. & DOHERTY, P. (1998) Neural cell adhesion molecule L1: relating disease to function. *Bioessays*, **20**, 668-75.

KENWRICK, S., WATKINS, A. & DE ANGELIS, E. (2000) Neural cell recognition molecule L1: relating biological complexity to human disease mutations. *Hum Mol Genet*, **9**, 879-86.

KISELYOV, V. V., BEREZIN, V., MAAR, T. E., SOROKA, V., EDVARDSEN, K., SCHOUSBOE, A. AND BOCK, E. (1997) The first immunoglobulin-like neural cell adhesion molecule (NCAM) domain is involved in double-reciprocal interaction with the second immunoglobulin-like NCAM domain and in heparin binding. *J Biol Chem*, **272**, 10125-34.

KLEYWEGT GJ & JONES TA (1997) Detecting folding motifs and similarities in protein structures. *Methods Enzymol*, **277**, 525-545.

KLEYWEGT, G., ZOU, J., KJELDGAARD, M. & AND JONES, T. (2001). Around O. M. ROSSMANN, E. ARNOLD (EDS.). *International Tables for Crystallography, Vol. F: Crystallography of*

Biological Macromolecules. pp. 353-356. Kluwer Academic Publishers Group, Dordrecht, The Netherlands.

KOSHLAND, D. E. (1958) Application of a theory of enzyme specificity tp protein synthesis. *PNAS*, **44**, 98-104.

KRAULIS, P. J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J Appl Cryst*, **24**, 946-950.

KUHN, T. B., STOECKLI, E. T., CONDRAU, M. A., RATHJEN, F. G. AND SONDEREGGER, P. (1991) Neurite outgrowth on immobilized axonin-1 is mediated by a heterophilic interaction with L1(G4). *J Cell Biol*, **115**, 1113-26.

KUNZ, B., LIERHEIMER, R., RADER, C., SPIRIG, M., ZIEGLER, U. AND SONDEREGGER, P. (2002) Axonin-1/TAG-1 mediates cell-cell adhesion by a cis-assisted trans-interaction. *J Biol Chem*, **277**, 4551-7.

KUNZ, S., SPIRIG, M., GINSBURG, C., BUCHSTALLER, A., BERGER, P., LANZ, R., RADER, C., VOGT, L., KUNZ, B. AND SONDEREGGER, P. (1998) Neurite fasciculation mediated by complexes of axonin-1 and Ng cell adhesion molecule. *J Cell Biol*, **143**, 1673-90.

LAEMMLI, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680-5.

LAI, E. C. (2004) Notch signaling: control of cell communication and cell fate. *Development*, **131**, 965-73.

LASKOWSKI, R. A., MACARTHUR, M. W., MOSS, D. S. & THORNTON, J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Cryst*, **26**, 283-291.

LASKOWSKI, R. A., MACARTHUR, M. W., MOSS, D. S. AND THORNTON, J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Cryst*, **26**, 283-291.

LAWRENCE, M. C. AND COLMAN, P.M. (1993) Shape complementarity at protein/protein interfaces. *J Mol Biol*, **234**, 946-50.

LEMMON, V., FARR, K. L. AND LAGENAUR, C. (1989) L1-mediated axon outgrowth occurs via a homophilic binding mechanism. *Neuron*, **2**, 1597-603.

LESK, A. M. & CHOTHIA, C. (1982) Evolution of proteins formed by beta-sheets. II. The core of the immunoglobulin domains. *J Mol Biol*, **160**, 325-42.

LUSTIG, M., SAKURAI, T. & GRUMET, M. (1999) Nr-CAM promotes neurite outgrowth from peripheral ganglia by a mechanism involving axonin-1 as a neuronal receptor. *Dev Biol*, **209**, 340-51.

LUSTIG, M., SAKURAI, T. AND GRUMET, M. (1999) Nr-CAM promotes neurite outgrowth from peripheral ganglia by a mechanism involving axonin-1 as a neuronal receptor. *Dev Biol*, **209**, 340-51.

MAEDA, Y., KOGA, H., YAMADA, H., UEDA, T. & IMOTO, T. (1995) Effective renaturation of reduced lysozyme by gentle removal of urea. *Protein Eng*, **8**, 201-5.

MALHOTRA, J. D., TSIOTRA, P., KARAGOGEOS, D. AND HORTSCH, M. (1998) Cis-activation of L1-mediated ankyrin recruitment by TAG-1 homophilic cell adhesion. *J Biol Chem*, **273**, 33354-9.

MARTINI, R. & SCHACHNER, M. (1986) Immunoelectron microscopic localization of neural cell adhesion molecules (L1, N-CAM, and MAG) and their shared carbohydrate epitope and myelin basic protein in developing sciatic nerve. *J Cell Biol*, **103**, 2439-48.

MATTEVI, A., TEDESCHI, G., BACCHELLA, L., CODA, A., NEGRI, A. & RONCHI, S. (1999) Structure of L-aspartate oxidase: implications for the succinate dehydrogenase/fumarate reductase oxidoreductase family. *Structure*, **7**, 745-56.

MATTEVI, A., VANONI, M. A., TODONE, F., RIZZI, M., TEPLYAKOV, A., CODA, A., BOLOGNESI, M. & CURTI, B. (1996) Crystal structure of D-amino acid oxidase: a case of active site mirror-image convergent evolution with flavocytochrome b2. *Proc Natl Acad Sci U S A*, **93**, 7496-501.

MERRITT, E. A. & BACON, D.J. (1997) Raster3D: Photorealistic Molecular Graphics. *Methods Enzymol*, **277**, 505-524.

MERRITT, E. A. AND BACON, D.J. (1997) Raster3D: Photorealistic Molecular Graphics. *Methods Enzymol*, **277**, 505-524.

MIELE, L. (2006) Notch signaling. *Clin Cancer Res*, **12**, 1074-9.

MILEV, P., MAUREL, P., HÄRING, M., MARGOLIS, R. K. AND MARGOLIS, R.U. (1996) TAG-1/axonin-1 is a high-affinity ligand of neurocan, phosphacan/protein-tyrosine phosphatase-zeta/beta, and N-CAM. *J Biol Chem*, **271**, 15716-23.

MIRANDA-RIOS, J., MORERA, C., TABOADA, H., DAVALOS, A., ENCARNACION, S., MORA, J. & SOBERON, M. (1997) Expression of thiamin biosynthetic genes (thiCOGE) and production of symbiotic terminal oxidase cbb3 in Rhizobium etli. *J Bacteriol*, **179**, 6887-93.

MOLLA, G., MOTTERAN, L., JOB, V., PILONE, M. S. & POLLEGIONI, L. (2003) Kinetic mechanisms of glycine oxidase from Bacillus subtilis. *Eur J Biochem*, **270**, 1474-82.

MOOS, M., TACKE, R., SCHERER, H., TEPLOW, D., FRÜH, K. & SCHACHNER, M. (1988) Neural adhesion molecule L1 as a member of the immunoglobulin superfamily with binding domains similar to fibronectin. *Nature*, **334**, 701-3.

MOOS, M., TACKE, R., SCHERER, H., TEPLOW, D., FRÜH, K. AND SCHACHNER, M. (1988) Neural adhesion molecule L1 as a member of the immunoglobulin superfamily with binding domains similar to fibronectin. *Nature*, **334**, 701-3.

MÖRTL, M. (2000) Klonierung und Expression der Liganden-Bindungsdomänen neuraler Zelladhasionsproteine der L1- und F11-Familie. *Universität Konstanz*, Dimplomarbeit.

MOSS, D. J. & WHITE, C.A. (1992) Solubility and posttranslational regulation of GP130/F11--a neuronal GPI-linked cell adhesion molecule enriched in the neuronal membrane skeleton. *Eur J Cell Biol*, **57**, 59-65.

MURSHUDOV, G. N., VAGIN, A. A. & DODSON, E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr*, **53**, 240-55.

NEW ENGLAND BIOLABS (2005) Taq DNA Polymerase. **Technical Bulletin #M0267(9/14/05)**, *New England Biolabs Inc.*, USA

NOVAGEN (2006) pET System Manual. **TB055 11th Edition 01/06**, *Merck KGaA*, Germany

OGAWA, J., KANEKO, H., MASUDA, T., NAGATA, S., HOSOYA, H. AND WATANABE, K. (1996) Novel neural adhesion molecules in the Contactin/F3 subgroup of the immunoglobulin superfamily: isolation and characterization of cDNAs from rat brain. *Neurosci Lett*, **218**, 173-6.

PAVLOU, O., THEODORAKIS, K., FALK, J., KUTSCHE, M., SCHACHNER, M., FAIVRE-SARRAILH, C. AND KARAGOGEOS, D. (2002) Analysis of interactions of the adhesion molecule TAG-1 and its domains with other immunoglobulin superfamily members. *Mol Cell Neurosci*, **20**, 367-81.

PAWELEK, P. D., CHEAH, J., COULOMBE, R., MACHEROUX, P., GHISLA, S. & VRIELINK, A. (2000) The structure of L-amino acid oxidase reveals the substrate trajectory into an enantiomerically conserved active site. *EMBO J*, **19**, 4204-15.

PERRIN, F. E., RATHJEN, F. G. & STOECKLI, E.T. (2001) Distinct subpopulations of sensory afferents require F11 or axonin-1 for growth to their target layers within the spinal cord of the chick. *Neuron*, **30**, 707-23.

PHILLPPSEN A (2001): DINO: Visualizing Structural Biology. Biozentrum, University of Basel, Basel, Switzerland (www.dino3d.org).

PILONE, M. S. (2000) D-Amino acid oxidase: new findings. *Cell Mol Life Sci*, **57**, 1732-47.

POLLEGIONI, L., CECILIANI, F., CURTI, B., RONCHI, S. & PILONE, M.S. (1995) Studies on the structural and functional aspects of Rhodotorula gracilis D-amino acid oxidase by limited trypsinolysis. *Biochem J*, **310 (Pt 2)**, 577-83.

PORTER, D. J., VOET, J. G. & BRIGHT, H.J. (1977) Mechanistic features of the D-amino acid oxidase reaction studied by double stopped flow spectrophotometry. *J Biol Chem*, **252**, 4464-73.

QIAGEN (2002) QIAquick Spin Handbook for QIAquick PCR Purification Kit, QIAquick Nucleotide Removal Kit, QIAquick Gel Extraction Kit. **July 2002**, *Qiagen GmbH*, Germany

QIAGEN (2003) The QIAexpressionist. **June 2003**, *Qiagen GmbH*, Germany

RADER, C., KUNZ, B., LIERHEIMER, R., GIGER, R. J., BERGER, P., TITTMANN, P., GROSS, H. AND SONDEREGGER, P. (1996) Implications for the domain arrangement of axonin-1 derived from the mapping of its NgCAM binding site. *EMBO J*, **15**, 2056-68.

RADER, C., STOECKLI, E. T., ZIEGLER, U., OSTERWALDER, T., KUNZ, B. AND SONDEREGGER, P. (1993) Cell-cell adhesion by homophilic interaction of the neuronal recognition molecule axonin-1. *Eur J Biochem*, **215**, 133-41.

RANSCHT, B. (1988) Sequence of contactin, a 130-kD glycoprotein concentrated in areas of interneuronal contact, defines a new member of the immunoglobulin supergene family in the nervous system. *J Cell Biol*, **107**, 1561-73.

RATHJEN, F. G. AND JESSELL, T.M. (1991) Glycoproteins that regulate the growth and guidance

of vertebrate axons: domains and dynamics of the immunoglobulin/fibronectin type III subfamily. *Semin Neurosci*, **3**, 297-307.

RATHJEN, F. G., WOLFF, J. M., FRANK, R., BONHOEFFER, F. & RUTISHAUSER, U. (1987) Membrane glycoproteins involved in neurite fasciculation. *J Cell Biol*, **104**, 343-53.

RICHARDSON, J. (1981) The anatomy and taxonomy of protein structures. *Advan Protein Chem*, **34**, 176-339.

RICKMAN, D. S., TYAGI, R., ZHU, X. X., BOBEK, M. P., SONG, S., BLAIVAS, M., MISEK, D. E., ISRAEL, M. A., KURNIT, D. M., ROSS, D. A., KISH, P. E. AND HANASH, S.M. (2001) The gene for the axonal cell adhesion molecule TAX-1 is amplified and aberrantly expressed in malignant gliomas. *Cancer Res*, **61**, 2162-8.

ROCHE DIAGNOSTICS (2004) Rapid DNA Ligation Kit. **Version December 2004**, *Roche Diagnostics GmbH*, Germany

ROSSMANN, M. G., MORAS, D. & OLSEN, K.W. (1974) Chemical and biological evolution of nucleotide-binding protein. *Nature*, **250**, 194-9.

SACCHI, S., LORENZI, S., MOLLA, G., PILONE, M. S., ROSSETTI, C. & POLLEGIONI, L. (2002) Engineering the substrate specificity of D-amino-acid oxidase. *J Biol Chem*, **277**, 27510-6.

SAHA, R., BAHADUR, R., PAL, A., MANDAL, S. AND CHAKRABARTI, P. (2006) ProFace: a server for the analysis of the physicochemical features of protein-protein interfaces. *BMC Struct Biol*, **6**, 11.

SAIER, M. H. J. & RAMSEIER, T.M. (1996) The catabolite repressor/activator (Cra) protein of enteric bacteria. *J Bacteriol*, **178**, 3411-7.

SAIER, M. H. J., CHAUVAUX, S., COOK, G. M., DEUTSCHER, J., PAULSEN, I. T., REIZER, J. & YE, J.J. (1996) Catabolite repression and inducer control in Gram-positive bacteria. *Microbiology*, **142 (Pt 2)**, 217-30.

SAMBROOK, J. & RUSSEL, D. (2001): Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

SCHÄGGER, H., CRAMER, W. A. & VON JAGOW, G. (1994) Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis. *Anal Biochem*, **217**, 220-30.

SCHREIER, H. (1993). Biosynthesis of glutamine and glutamate and the assimilation of ammonia. A. SONENSHEIN, J. HOCH & R. LOSICK (EDS.). *Bacillus subtilis and Other Gram-positive Bacteria: Biochemistry, Physiology, and Molecular Biology*. pp. 281-298. American Society for Microbiology, Washington, D. C.

SCHREUDER, H. A., PRICK, P. A., WIERENGA, R. K., VRIEND, G., WILSON, K. S., HOL, W. G. & DRENTH, J. (1989) Crystal structure of the p-hydroxybenzoate hydroxylase-substrate complex refined at 1.9 A resolution. Analysis of the enzyme-substrate and enzyme-product complexes. *J Mol Biol*, **208**, 679-96.

SCHÜRMANN, G., HASPEL, J., GRUMET, M. AND ERICKSON, H.P. (2001) Cell adhesion molecule L1 in folded (horseshoe) and extended conformations. *Mol Biol Cell*, **12**, 1765-73.

SETTEMBRE, E. C., DORRESTEIN, P. C., PARK, J., AUGUSTINE, A. M., BEGLEY, T. P. & EALICK, S.E. (2003) Structural and mechanistic studies on ThiO, a glycine oxidase essential for thiamin biosynthesis in Bacillus subtilis. *Biochemistry*, **42**, 2971-81.

SETTEMBRE, E. C., DORRESTEIN, P. C., PARK, J., AUGUSTINE, A. M., BEGLEY, T. P. AND EALICK, S.E. (2003) Structural and mechanistic studies on ThiO, a glycine oxidase essential for thiamin biosynthesis in Bacillus subtilis. *Biochemistry*, **42**, 2971-81.

SHAPIRO, L., FANNON, A. M., KWONG, P. D., THOMPSON, A., LEHMANN, M. S., GRÜBEL, G., LEGRAND, J. F., ALS-NIELSEN, J., COLMAN, D. R. AND HENDRICKSON, W.A. (1995) Structural basis of cell-cell adhesion by cadherins. *Nature*, **374**, 327-37.

SIMONATI, A., BOARETTO, F., VETTORI, A., DABRILLI, P., CRISCUOLO, L., RIZZUTO, N. & MOSTACCIUOLO, M.L. (2006) A novel missense mutation in the L1CAM gene in a boy with L1 disease. *Neurol Sci*, **27**, 114-7.

SOROKA, V., KOLKOVA, K., KASTRUP, J. S., DIEDERICHS, K., BREED, J., KISELYOV, V. V., POULSEN, F. M., LARSEN, I. K., WELTE, W., BEREZIN, V., BOCK, E. AND KASPER, C. (2003) Structure and interactions of NCAM Ig1-2-3 suggest a novel zipper mechanism for homophilic adhesion. *Structure*, **11**, 1291-301.

STOECKLI, E. T. & LANDMESSER, L.T. (1995) Axonin-1, Nr-CAM, and Ng-CAM play different roles in the in vivo guidance of chick commissural neurons. *Neuron*, **14**, 1165-79.

STOECKLI, E. T. AND LANDMESSER, L.T. (1995) Axonin-1, Nr-CAM, and Ng-CAM play different roles in the in vivo guidance of chick commissural neurons. *Neuron*, **14**, 1165-79.

STOECKLI, E. T., KUHN, T. B., DUC, C. O., RUEGG, M. A. AND SONDEREGGER, P. (1991) The axonally secreted protein axonin-1 is a potent substratum for neurite growth. *J Cell Biol*, **112**, 449-55.

STOECKLI, E. T., LEMKIN, P. F., KUHN, T. B., RUEGG, M. A., HELLER, M. AND SONDEREGGER, P. (1989) Identification of proteins secreted from axons of embryonic dorsal-root-ganglia neurons. *Eur J Biochem*, **180**, 249-58.

STOECKLI, E. T., SONDEREGGER, P., POLLERBERG, G. E. & LANDMESSER, L.T. (1997) Interference with axonin-1 and NrCAM interactions unmasks a floor-plate activity inhibitory for commissural axons. *Neuron*, **18**, 209-21.

STOECKLI, E. T., ZIEGLER, U., BLEIKER, A. J., GROSCURTH, P. AND SONDEREGGER, P. (1996) Clustering and functional cooperation of Ng-CAM and axonin-1 in the substratum-contact area of growth cones. *Dev Biol*, **177**, 15-29.

STRATAGENE (2005) BL21-CodonPlus Competent Cells. **Revision #085008**, *Stratagene Inc.*, USA

STRATAGENE (2006) QuikChange Multi Site-Directed Mutagenesis Kit. **Revision #066005b**, *Stratagene Inc.*, USA

SU, X. D., GASTINEL, L. N., VAUGHN, D. E., FAYE, I., POON, P. AND BJORKMAN, P.J. (1998) Crystal structure of hemolin: a horseshoe shape with implications for homophilic adhesion. *Science*, **281**, 991-5.

SUTER, D. M., POLLERBERG, G. E., BUCHSTALLER, A., GIGER, R. J., DREYER, W. J. AND

SONDEREGGER, P. (1995) Binding between the neural cell adhesion molecules axonin-1 and Nr-CAM/Bravo is involved in neuron-glia interaction. *J Cell Biol*, **131**, 1067-81.

TANG, N., HE, M., O'RIORDAN, M. A., FARKAS, C., BUCK, K., LEMMON, V. & BEARER, C.F. (2006) Ethanol inhibits L1 cell adhesion molecule activation of mitogen-activated protein kinases. *J Neurochem*, **96**, 1480-90.

TERWILLIGER, T. C. (2000) Maximum-likelihood density modification. *Acta Crystallogr D Biol Crystallogr*, **56**, 965-72.

TERWILLIGER, T. C. & BERENDZEN, J. (1999) Evaluation of macromolecular electron-density map quality using the correlation of local r.m.s. density. *Acta Crystallogr D Biol Crystallogr*, **55**, 1872-7.

TESSIER-LAVIGNE, M. AND GOODMAN, C.S. (1996) The molecular biology of axon guidance. *Science*, **274**, 1123-33.

TRAKA, M., DUPREE, J. L., POPKO, B. AND KARAGOGEOS, D. (2002) The neuronal adhesion protein TAG-1 is expressed by Schwann cells and oligodendrocytes and is localized to the juxtaparanodal region of myelinated fibers. *J Neurosci*, **22**, 3016-24.

TRICKEY, P., WAGNER, M. A., JORNS, M. S. & MATHEWS, F.S. (1999) Monomeric sarcosine oxidase: structure of a covalently flavinylated amine oxidizing enzyme. *Structure*, **7**, 331-45.

TSIOTRA, P. C., KARAGOGEOS, D., THEODORAKIS, K., MICHAELIDIS, T. M., MODI, W. S., FURLEY, A. J., JESSELL, T. M. & PAPAMATHEAKIS, J. (1993) Isolation of the cDNA and chromosomal localization of the gene (TAX1) encoding the human axonal glycoprotein TAG-1. *Genomics*, **18**, 562-7.

TSIOTRA, P. C., KARAGOGEOS, D., THEODORAKIS, K., MICHAELIDIS, T. M., MODI, W. S., FURLEY, A. J., JESSELL, T. M. AND PAPAMATHEAKIS, J. (1993) Isolation of the cDNA and chromosomal localization of the gene (TAX1) encoding the human axonal glycoprotein TAG-1. *Genomics*, **18**, 562-7.

TSIOTRA, P. C., THEODORAKIS, K., PAPAMATHEAKIS, J. AND KARAGOGEOS, D. (1996) The fibronectin domains of the neural adhesion molecule TAX-1 are necessary and sufficient for homophilic binding. *J Biol Chem*, **271**, 29216-22.

UMHAU, S., POLLEGIONI, L., MOLLA, G., DIEDERICHS, K., WELTE, W., PILONE, M. S. & GHISLA, S. (2000) The x-ray structure of D-amino acid oxidase at very high resolution identifies the chemical mechanism of flavin-dependent substrate dehydrogenation. *Proc Natl Acad Sci U S A*, **97**, 12463-8.

VAGIN, A. & TEPLYAKOV, A. (1997) MOLREP: an Automated Program for Molecular Replacement. *J Appl Cryst*, **30**, 1022-5.

VAGIN, A. AND TEPLYAKOV, A. (1997) MOLREP: an Automated Program for Molecular Replacement. *J Appl Cryst*, **30**, 1022-5.

VAN DER HORN, P. B., BACKSTROM, A. D., STEWART, V. & BEGLEY, T.P. (1993) Structural genes for thiamine biosynthetic enzymes (thiCEFGH) in Escherichia coli K-12. *J Bacteriol*, **175**, 982-92.

VOGT, L., SCHRIMPF, S. P., MESKENAITE, V., FRISCHKNECHT, R., KINTER, J., LEONE, D. P., ZIEGLER, U. & SONDEREGGER, P. (2001) Calsyntenin-1, a proteolytically processed postsynaptic membrane protein with a cytoplasmic calcium-binding domain. *Mol Cell Neurosci*, **17**, 151-66.

VOGT, L., SCHRIMPF, S. P., MESKENAITE, V., FRISCHKNECHT, R., KINTER, J., LEONE, D. P., ZIEGLER, U. AND SONDEREGGER, P. (2001) Calsyntenin-1, a proteolytically processed postsynaptic membrane protein with a cytoplasmic calcium-binding domain. *Mol Cell Neurosci*, **17**, 151-66.

WAGNER, M. A. & JORNS, M.S. (2000) Monomeric sarcosine oxidase: 2. Kinetic studies with sarcosine, alternate substrates, and a substrate analogue. *Biochemistry*, **39**, 8825-9.

WAGNER, M. A. & SCHUMAN JORNS, M. (1997) Folate utilization by monomeric versus heterotetrameric sarcosine oxidases. *Arch Biochem Biophys*, **342**, 176-81.

WEINER, L. M. (1999) An overview of monoclonal antibody therapy of cancer. *Semin Oncol*, **26**, 41-50.

WIERENGA, R. K., DRENTH, J. & SCHULZ, G.E. (1983) Comparison of the three-dimensional protein and nucleotide structure of the FAD-binding domain of p-hydroxybenzoate hydroxylase with the FAD- as well as NADPH-binding domains of glutathione reductase. *J Mol Biol*, **167**, 725-39.

WILKEMEYER, M. F., CHEN, S., MENKARI, C. E., BRENNEMAN, D. E., SULIK, K. K. & CHARNESS, M.E. (2003) Differential effects of ethanol antagonism and neuroprotection in peptide fragment NAPVSIPQ prevention of ethanol-induced developmental toxicity. *Proc Natl Acad Sci U S A*, **100**, 8543-8.

WINN, M. D., ISUPOV, M. N. AND MURSHUDOV, G.N. (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr D Biol Crystallogr*, **57**, 122-33.

WOOD, P. M., SCHACHNER, M. & BUNGE, R.P. (1990) Inhibition of Schwann cell myelination in vitro by antibody to the L1 adhesion molecule. *J Neurosci*, **10**, 3635-45.

YAMAMOTO, M., BOYER, A. M., CRANDALL, J. E., EDWARDS, M. AND TANAKA, H. (1986) Distribution of stage-specific neurite-associated proteins in the developing murine nervous system recognized by a monoclonal antibody. *J Neurosci*, **6**, 3576-94.

YOSHIHARA, Y., KAWASAKI, M., TAMADA, A., NAGATA, S., KAGAMIYAMA, H. AND MORI, K. (1995) Overlapping and differential expression of BIG-2, BIG-1, TAG-1, and F3: four members of an axon-associated cell adhesion molecule subgroup of the immunoglobulin superfamily. *J Neurobiol*, **28**, 51-69.

ZHAO, G. & JORNS, M.S. (2002) Monomeric sarcosine oxidase: evidence for an ionizable group in the E.S complex. *Biochemistry*, **41**, 9747-50.

ZUELLIG, R. A., RADER, C., SCHROEDER, A., KALOUSEK, M. B., VON BOHLEN UND HALBACH, F., OSTERWALDER, T., INAN, C., STOECKLI, E. T., AFFOLTER, H. U. & FRITZ, A. (1992) The axonally secreted cell adhesion molecule, axonin-1. Primary structure, immunoglobulin-like and fibronectin-type-III-like domains and glycosyl-phosphatidylinositol anchorage. *Eur J Biochem*, **204**, 453-63.