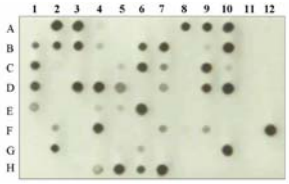
Fusion proteins and fusion tags

Most frequently used tags are His-6 tags for IMAC (Immobilized metal-ion affinity) capture

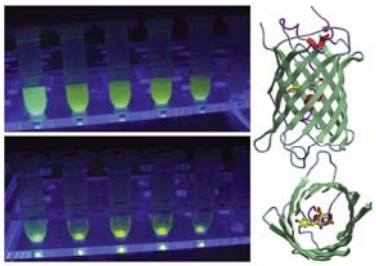
Figure 4-12 Generic design of a tagged fusion construct. The protein construct contains an affinity tag for easy affinity capture with a shorter linker to a solubility enhancing fusion protein finally linked to the target sequence. The cleavage site is located right before the N-terminus of the target sequence to leave as few residues as possible behind. The Gateway p-DEST-HisMBP vector is a typical example,[9] with the TEV cleavage site providing for simple IMAC purification when a His-tagged TEV protease is used (Figure 4-13).

Figure 4-13 His-tag antibody Western blot of C-terminally His_6-tagged proteins. A stained immuno-blot of lysates from a small scale, parallel expression experiment in 96-well format allows rapid identification of colonies that have overexpressed a large amount of full length constructs (intense dark spots). Figure reproduced[10] with permission from Springer Verlag.



His-tagged GFP fusion and His6-TEV

Figure 4-14 Green fluorescent protein (GFP). GFP is a spontaneously fluorescent 238 amino acid protein isolated from the Pacific jellyfish, *Aequoria victoria*.[11] Its biological function is to transduce via energy transfer the blue chemiluminescence of another jellyfish protein, aequorin, into green fluorescent light. The fluorophore originates from the internal S66-Y67-G68 sequence (shown inside the β-barrel in ball-and-stick presentation) which is posttranslationally modified to a 4-(p-hydroxybenzylidene)imidazolidin-5-one chromophore structure. The images of the test tubes show both supernatant and cell pellets (bottom rack) containing expressed GFP-fusion protein at various levels. The rightmost tube reveals no expression at all. Note that the leftmost sample contains an engineered cyan fluorescing construct of GFP. Image of fluorescent samples courtesy of Gunter Stier, MPI Heidelberg, Germany.



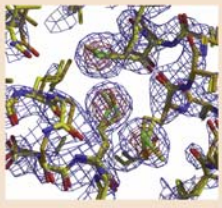
His-tagged proteins where the tag is cleaved off by a His-tagged TEV protease allow fast 2-step IMAC purification

Se-Met labeling of proteins

Se-Met labeling - largely by metabolic starvation techniques - produces isomorphous structures with anomalous marker atoms for anomalous phasing.

				dist(Å)	σU(Å)
MET	CG	SD	covalent	1.803	0.025
MSE	CG	SE	covalent	1.950	0.030
MET	SD	CE	covalent	1.791	0.025
MSE	SE	CE	covalent	1.950	0.030
				angle(°)	σU(°)
MSE	CG	SE	CE	98.923	2.250
MET	CG	SD	CE	100.900	2.250

Table 4-1 Bond lengths and bond angles for Se-Met and native S-Met.



Box 4-4 Selenomethionine labeling. Incorporation of Se-Met (instead of regular methionine residues provides heavy atom labels for anomalous phasing while maintaining structural isomorphism to the native protein. One method of incorporating Se-Met is by expression in a methionine auxotroph host strain and supplementing the minimal medium with Se-Met. In the metabolic inhibition technique bacteria are grown in regular medium and Met biosynthesis inhibiting amino acids Ile, Lys, and Thr as well as Se-Met are added to the medium before induction.

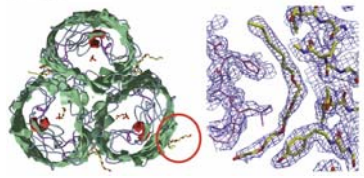
Figure 4-15 Isomorphism between selenomethionine mutants and native methionine crystal structures. The figure shows a superposition of a Se-Met structure (yellow sticks, in its electron density) and its native counterpart (orange sticks). Shown is the crystal structure core containing three Met residues of the apolipoprotein E3 (ApoE3) 22 kDa N-terminal domain. The situation represents a worst-case scenario. There are three Met in close proximity; the independently refined Se-Met structure crystallizes in a different crystal form (probably because of absence of 2-β-mercaptoethanol (2BME) in the crystallization cocktail), and the structure itself is a flexible helix bundle. Nonetheless, the local coordinate r.m.s. deviations for the Met residues (0.4 Å) are on the order of coordinate uncertainty of the individual structures (0.28 and 0.44 Å respectively). PDB entries 1bz4 and 1orz.[12]

Membrane protein expression

Hydrophobic transmembrane stem needs protection

Detergents are added for solubilization and crystallization

Figure 4-18 The trimeric bacterial outer membrane protein OprP. Published in 2007, the OprP structure revealed a new structural motif in bacterial membrane proteins, a 9-residue arginine ladder. OprP controls the transport of essential phosphate anions into the pathogenic bacterium *Pseudomonas aeruginosa*. The Arg ladder extends from the extracellular surface down to a constriction zone where a phosphate anion is bound (visible in the center of the channels). Note that the crystal structure again contains detergent molecules along the transmembrane barrel, in this case hydroxyethyloxytriethyloxyloctane or C8E. The right panel shows the region around the circled detergent molecule in electron density. In this position, the detergent in addition mediates a crystal contact to a symmetry related molecule (magenta). PDB entry 2o4v.[13]



Fusions for transport and membrane insertion

Sidebar 4-8 Mystic Mistic. The recent discovery of Mistic (Membrane Integrating Sequence for Translation of Integral membrane protein Constructs) has opened another promising avenue for membrane protein expression. Mistic is an unusual *Bacillus subtilis* dual-topology protein that folds and inserts autoenomously into the membrane forming an integral four-helix bundle membrane protein (Figure 4-16).[108] In addition, Mistic lacks any signal sequence and bypasses the cellular translocon complex, and therefore provides a highly suitable fusion partner for expression of eukaryotic membrane proteins.

The auto-inserting ability of Mistic together with the bypassing of the cellular translocon has raised the probability that more structures of crucial membrane proteins of therapeutic interest, such as ion channels and 7-TM G-protein coupled receptors (GPCRs), will become available soon. More than 1000 GPCRs exist, which are involved in regulation of a wide variety of biological functions, and many of them are drug targets. Another avenue to expression of GPCRs via antibody scaffolding[109] was outlined in Chapter 3, and the scaffolding via the insertion of a T4-lysozyme mutant into a flexible loop of the GPCR structure.[110]

Similar to the idea of using Mistic as a transport and integration fusion, members of the bacterial Tol proteins[111] export a wide range of molecules across the periplasmic space and outer membrane of Gram-negative bacteria, and might provide a further avenue for membrane protein expression in bacteria. An engineered version of the anti-apoptotic Bcl-2 family protein Bcl-XL has also been used in a fusion system for membrane proteins.[112]

A good resource for membrane protein production and structure is Stevens White's web page: http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html

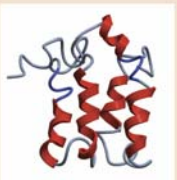


Figure 4-16 NMR solution structure of Mistic. The ribbon diagram shows the lowest energy conformer of the irregular four-helix bundle of Mistic. The lipid facing surfaces are unusually polar, and the presence of a concentric ring of solubilizing LDAO detergent molecules was deduced from NMR NOE interactions. PDB entry 1ygm.[108]

Chemical modification by surface Lys methylation

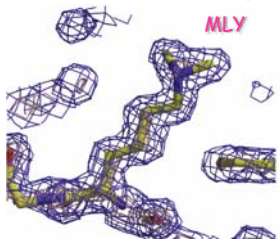


Figure 4-19 Chemically modified lysine residue (dimethyllysine). The electron density clearly shows that this surface residue of the alanine racemase of *Bacillus anthracis* is dimethylated (MLY). See Figure 2-27 for the density of a natively occurring, posttranslationally modified trimethyllysine (M3L) in calmodulin. PDB entry 2vd8.[107]

Lysine methylation - LYS to MLY - is largely used as a salvage strategy and also reduces solubility and changes charge distribution - multiple effects and actual causality for success is hard to establish