Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Expression, purification and crystallization of the cofactor-independent monooxygenase SnoaB from the nogalamycin biosynthetic pathway

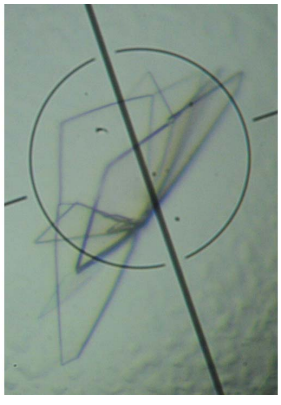
Hanna Koskiniemi,[a]‡ Thadee Grocholski,[b]‡ Gunter Schneider[a]* and Jarmo Niemi[b]*

[a]Department of Medical Biochemistry and Biophysics, Karolinska Institutet, S-171 77 Stockholm, Sweden, and [b]Department of Biochemistry and Food Chemistry, University of Turku, FIN-20014 Turku, Finland

‡ These authors contributed equally.

Correspondence e-mail: gunter.schneider@ki.se, jarnie@utu.fi

Received 23 December 2008
Accepted 12 January 2009



12-deoxy-nogalonic acid oxygenase (SnoaB) catalyzes the oxygenation of 12-deoxy-nogalonic acid at position 12 to yield nogalonic acid, which is one of the steps in the biosynthesis of the polyketide nogalamycin in *Streptomyces nogalater*. SnoaB belongs to a family of small cofactor-free oxygenases which carry out oxygenation reactions without the aid of any prosthetic group, cofactor or metal ion. Recombinant SnoaB was crystallized in space group $P2_12_12$, with unit-cell parameters a = 58.8, b = 114.1, c = 49.5 Å, and these crystals diffracted to 2.4 Å resolution. Recombinant SnoaB does not contain any methionine residues and three double mutants were designed and produced for the preparation of selenomethionine-substituted samples. The selenomethionine-substituted mutant F40M/L89M crystallized in the same space group as the native enzyme.

1. Introduction

Anthracycline antibiotics are a class of aromatic polyketides that are mainly produced by Gram-positive bacteria of the genus *Streptomyces*. They are of great significance in medicine because of their cytostatic properties: some of these compounds, for instance doxorubicin and epirubicin, are amongst the most commonly used anticancer agents worldwide (for a recent review, see Minotti *et al.*, 2004). The biosynthesis of aromatic polyketides begins with assembly of the aglycone scaffold, which in the case of anthracyclines is the 7,8,9,10-tetrahydrotetracene-5,12-quinone core. The carbon backbone of the aglycone is assembled from acetate and malonate units by an iterative polyketide synthetase type II (PKSII) multienzyme complex and is followed by cyclization/reduction reactions catalyzed by aromatases, ketoreductases and cyclases. The aglycone is subsequently modified by various tailoring enzymes, including methylases, hydroxylases and aminotransferases. One or several sugar moieties are attached to the aglycone skeleton during the tailoring stage to convert these metabolites into their biologically active derivatives (for recent reviews, see Schneider, 2005; Fischbach & Walsh, 2006; Hertweck *et al.*, 2007; Metsä-Ketelä *et al.*, 2008).

SnoaB is a tailoring enzyme that acts in the biosynthetic pathway of the anthracycline antibiotic nogalamycin in *S. nogalater*. It catalyzes the oxygenation at position 12 of the aglycone prior to the closure of the fourth ring (Fig. 1). This reaction yields the first isolatable intermediate of nogalamycin biosynthesis, nogalonic acid (Ylihonko, Tuikkanen *et al.*, 1996). The *snoaB* gene encodes a small

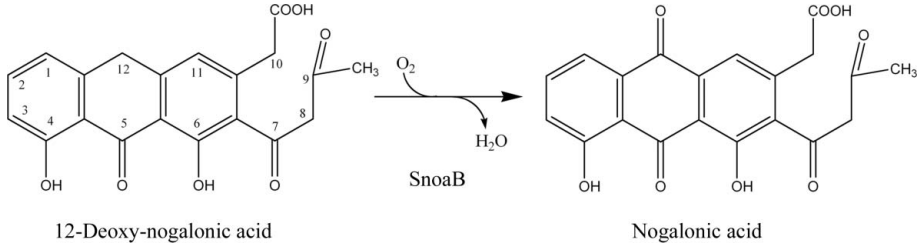


Figure 1
Reaction catalyzed by SnoaB.

protein of 118 amino-acid residues and sequence comparisons suggest that SnoaB might be a member of a family of bacterial monooxygenases that carry out oxygenation reactions in the absence of any prosthetic group, cofactor or metal ion. Although several three-dimensional structures of members of this sequence family are available, only two have confirmed monooxygenase activity: ActVA-Orf6 (Sciara *et al.*, 2003) and YgiN (Adams & Jia, 2005). ActVA-Orf6 is a tailoring enzyme that is involved in the actinorhodin biosynthetic pathway in *S. coelicolor* strain A3(2), but it is only distantly related to SnoaB (<18% sequence identity), although its enzymatic function is similar.

The intriguing ability of these monooxygenases to catalyze the incorporation of oxygen derived from molecular oxygen in the absence of any spin-delocalizing devices such as metal ions or cofactors has until recently been relatively poorly understood in mechanistic terms (Fetzner, 2002, 2007). Initial structural insights have been obtained from studies of ActVA-Orf6 (Sciara *et al.*, 2003), but a number of mechanistic issues remain unsolved. One question in particular deserves attention, namely whether or not the mechanism of oxygen activation proposed for ActVA-Orf6 holds for the whole sequence family and is conserved in other structurally unrelated cofactor-independent oxygenases (Beinker *et al.*, 2006; Widboom *et al.*, 2007; Stec & Stieglitz, 2008). As part of our studies of the enzymology of aromatic polyketide biosynthesis, we here report the cloning, recombinant expression and crystallization of SnoaB from *S. nogalater*.

2. Materials and methods

2.1. Cloning

The gene coding for SnoaB was amplified by PCR from the plasmid template pSn15 (Ylihonko, Hakala *et al.*, 1996) using Phusion DNA polymerase (Finnzymes). The oligonucleotides used for amplification were 5′-TCG **AGA TCT** CCG ACA CGA GTG AAC GAT GGC GTC-3′ (*Bgl*II restriction site in bold) and 5′-GCC **AAG CTT** CAG CGA TGT CCG GAG CCG GTC G-3′ (*Hin*dIII restriction site in bold). The PCR product was then digested and ligated into a pBAD/HisB vector (Invitrogen) modified as described previously (Kallio *et al.*, 2006). The construct encodes a protein with an N-terminal polyhistidine tag (MAHHHHHHHRS) with the first residue (methionine) of the native sequence excluded. The resulting plasmid was transformed into *Escherichia coli* strain TOP10 (Invitrogen) and the DNA sequence was verified by sequencing three independent clones of the expression plasmid.

2.2. Design and cloning of the mutations

With the exception of the N-terminal residue, the amino-acid sequence of SnoaB does not contain any methionines. As the recombinant SnoaB protein consists of only 128 residues, two selenomethionines are likely to provide sufficient phasing power for structure determination. Therefore, three different double mutants were designed. Mutation sites were selected such that the methionines would be located in predicted α-helices in order to minimize their flexibility. For the same reason, the replacement of larger hydrophobic amino acids from the hydrophobic protein interior was preferred (Leahy *et al.*, 1994; Gassner & Matthews, 1999). Based on secondary-structure predictions and sequence alignments with homologous proteins, the residues Phe29, Phe40 and Leu89 were selected. Among homologous enzymes of *Streptomyces* origin, these three positions are conserved as hydrophobic residues (Fig. 2). Furthermore, a methionine is present at the position of Phe40 in several of these enzymes. The three residues are all located in helical regions according to secondary-structure predictions with the programs *JPred*2 (Cuff & Barton, 2000) and *GOR*4 (Combet *et al.*, 2000). Phe29, Phe40 and Leu89 were replaced by methionines in pairs to produce three different double mutants. The mutations (F29M/F40M, F29M/L89M and F40M/L89M) were introduced into the expression vector by the four-primer method described by Higuchi *et al.* (1988). All constructs were verified by DNA sequencing.

2.3. Expression and purification

The wild-type and mutant constructs of SnoaB were expressed in *E. coli* strain TOP10 in either 5 l fermentation or 2 l shaker-flask cultures at 303 K in 2×YT medium. Gene expression was induced with L-arabinose to a final concentration of 0.02%(*w*/*v*) at an OD_{600} of 0.4–0.6. Expression cultures were grown at 303 K for 16 h. Selenomethionine-substituted SnoaB mutants were produced using a previously described protocol (Sultana *et al.*, 2007).

Purification of SnoaB was performed on an ÄKTA FPLC system at 277–281 K using an Ni^{2+}-affinity column (5 ml HisTrap HP) and a gel-filtration column (HiLoad 26/10 Superdex 200 prep-grade, GE Healthcare). The Ni^{2+}-affinity column was equilibrated with five column volumes of buffer *A* (0.02 *M* Na_3PO_4, 0.5 *M* NaCl, 40 m*M* imidazole pH 7.6), five column volumes of buffer *B* (0.02 *M* Na_3PO_4,

```
snoab       ------MPTR----VNDGVDADEVTFVNRFTVHGAPAEFESVFARTAAFFARQPG-----FVRHTLLR-- 53
AknX        ------MTDH----EPGTEGADAVTFVNTFTVHAEPEVFEKEFARTSEFMAQQPG-----FVRHTLCR-- 53
Stf0I       ------MPTH----KYDATDSGIVTFVNQFTVHSSPEEFEKIFAEVSEFMAEQPG-----FIQYTLSRSI 55
CosX        MRSTDPMSAD----QPGPASGGPATFVNSFTLRTTPEEFEDVFARTARFMERQPG-----FLGYTLVR-- 59
Gris        ------MDR-------NGPPPGAVTFINRFTVTGDPEEFEAAFAKVAAFMTARPG-----ILGHTLSR-- 50
DauA        ------MEQRCLWPEPNDAGSGSVTFVNRFTLSGSAEDFEAAFAETAEFLCRRPGFRWHVLLAPTGSG-- 62
ChaH        ------MREK-----------PMLVFINRFTVSGDPLEFERAFAETAAHLARQPG-----FIRYRLVR-- 46

snoab       ERDKDNSYVNIAVWTDHDAFRRALAQPGFLPHATALRALSTSEHGLFTA---RQTLPEGGDTTGSGHR- 118
AknX        HAERPGQYVNVAEWRDLASFRAAVSHDGFRPHADALRALSESRPELYLVRLRREGAPGLDGPASEGEEI 122
Stf0I       DEDKQDRYINIALWEDAQSWRNAVAHPGFQDHAKEIRARTTNVGELYAP---RQSFSVK---------- 111
CosX        HLEQPHSYVNIARWADVASFRAAVGQSDFRPHAEALRAISTSSSNLYLE--RR---------SATGEAR 117
Gris        HLDEPGQYVNVAVWRDAGSLRAAVAHPDFGAHAGELRRLATSESDVYVERQRHLG-GAGGR---QGEAV 115
DauA        SADVRPQYVNIAVWDDEASFRAAVAHPQFPAHAAVLRALSTSEPTLYRSRQIRVAPGAPAMSRPEGRTT 131
ChaH        STEDPHSYANVAEWETAAALREALRGPEFDDHARRLRELARSEPQFYDVVTETDAIRSRR--------- 106
```

Figure 2
Sequence alignment of SnoaB with homologous enzymes from various *Streptomyces* strains. The residues mutated to methionines in SnoaB in this study are highlighted in grey. AknX denotes an oxygenase from *S. galilaeus* (GenBank accession code AAF70105); Stf0I is from *S. steffisburgensis* (CAJ42321), CosX from *S. olindensis* (ABC00737), Gris from *S. griseoruber* (AAP85340), DauA from *Streptomyces* sp. C5 (AAA87617) and ChaH from *S. chartreusis* (CAH10167).

0.5 *M* NaCl, 0.5 *M* imidazole pH 7.6) and again with buffer *A*. The harvested cells were broken with a French press in buffer *A* containing 1 m*M* phenylmethanesulfonyl fluoride and 1 m*M* benzamidine and the cell-free lysate was applied onto the column. A standard run included, after injection, a 100–150 ml wash (buffer *A*), a linear gradient (40–50 ml) to 100% buffer *B* and 20–50 ml of 100% buffer *B*. The peak eluting at 15–20 ml after the start of the gradient was concentrated (Millipore Centriprep YM-10) to 5–10 ml and applied onto the gel-filtration column. The sample eluted at 230–250 ml from a pre-equilibrated (100 m*M* Tris–HCl, 150 m*M* NaCl pH 8.0) column. For the selenomethionine SnoaB mutant, 1.5 m*M* DTT was included in the gel-filtration running buffer.

The purity of the protein was verified by SDS–PAGE. The incorporation of selenomethionine was confirmed by electrospray mass spectrometry (ES–MS). The protein was stored until further use in 50 m*M* Tris–HCl pH 8.0, 75 m*M* NaCl, 50%(*v*/*v*) glycerol at 253 or 193 K.

2.4. Quaternary structure verification

Cross-linking with glutaraldehyde and electrophoretic analysis were carried out as described previously (Jämsen *et al.*, 2007) with minor modifications: 64 µ*M* (1 mg ml^{-1}) SnoaB in 1× PBS buffer pH 7.4 was incubated with 260 m*M* glutaraldehyde at 297 K for 60 min.

2.5. Assay of SnoaB oxygenase activity

Danthron (Sigma) was reduced to a substrate analogue, dithranol, as described previously (Cross & Perkin, 1930). The oxygenase activity of SnoaB was followed by observing the formation of danthron from dithranol spectrophotometrically at 430 nm. The reaction mixture consisted of 103 µ*M* dithranol and 40%(*v*/*v*) ethylene glycol in 57 m*M* TES buffer pH 7.0 in a final volume of 600 µl. The reaction was started by the addition of 10 µl SnoaB at 5 mg ml^{-1} (final concentration 5.8 µ*M*).

2.6. Crystallization

The protein dissolved in the storage buffer was diluted tenfold with 50 m*M* Tris–HCl pH 8.0, 50 m*M* NaCl. The enzyme was incubated with an excess of the product analogue danthrone (saturated in acetonitrile) for 1 h on ice at a volume ratio of one part ligand to seven parts protein solution. The protein was then concentrated with Microsep and Nanosep 3K Omega centrifugal concentrator devices (Pall Corporation) and was filtered with Ultrafree MC 0.1 µm filters (Millipore) prior to setting up the crystallization experiments. The protein concentration used in these experiments was in the range 35–40 mg ml^{-1} as determined using the Bradford assay. Crystallization screening was carried out in 96-well plates with sitting drops and a Phenix crystallization robot using various sparse-matrix screens from Qiagen and Hampton Research. The drops were set up with 70 µl reservoir-solution volume and a drop size of 0.4 µl (0.2 µl protein solution and 0.2 µl reservoir solution). Index Screen (Hampton Research) resulted in four crystal hits, but the crystals were not reproducible and/or of poor diffraction quality. Diffraction-quality crystals were finally obtained from the NeXtal JCSG+ sparse-matrix screen (Qiagen) at 277 K in 0.4 µl drops with 80 µl reservoir-solution volume. This condition was further optimized by varying the precipitant concentration and by using a commercial additive screen (Hampton Research) with either the sitting-drop or hanging-drop vapour-diffusion method. For crystal optimization, the drop size was typically increased to 2 µl and the volume of the reservoir solution was 1.0 ml.

The selenomethionine-substituted mutant was crystallized in a similar manner to the wild-type protein, but streak-seeding from native crystals was employed after ∼16 h of equilibration at 277 K. The protein concentration in these experiments was in the range 20–25 mg ml^{-1}.

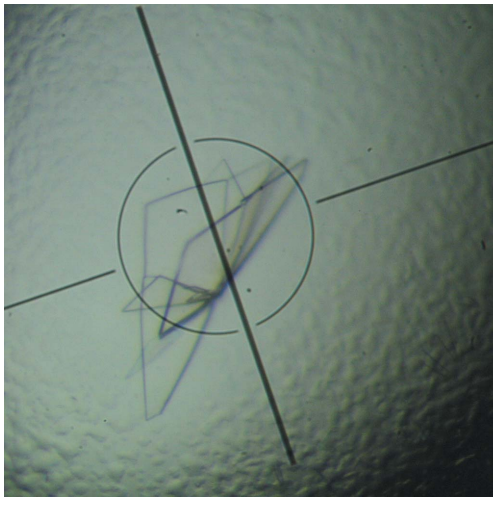


Figure 3
Crystals of the selenomethionine-substituted F40M/L89M double mutant of SnoaB.

2.7. Data collection and analysis

Crystals were cryoprotected by rapid soaking in reservoir solution containing 25%(*v*/*v*) ethylene glycol and then flash-frozen in liquid nitrogen. A data set was collected to a resolution of 2.4 Å from a wild-type crystal on beamline ID23-2 at ESRF. The data set was processed with the program *MOSFLM* (Leslie, 1992) and scaled with *SCALA* from the *CCP*4 package (Collaborative Computational Project, Number 4, 1994).

3. Results and discussion

The gene coding for SnoaB was cloned in a modified pBAD-HisB expression vector. Sequencing of three plasmids containing the *snoaB* gene from three separate PCR reactions revealed a discrepancy at position 25 in the amino-acid sequence. The sequences uniformly code for an alanine (codon GCG), while the sequence in GenBank (gi:2916810) codes for a glycine (codon GGG), suggesting a possible error in the deposited sequence.

SnoaB with an N-terminal His_7 tag was purified to homogeneity using a two-step purification protocol. The typical yield of pure protein ranged from 5 to 40 mg per litre of culture medium.

Cross-linking of SnoaB with glutaraldehyde indicated that it is present as a dimer in solution (data not shown). Three methionines were introduced into the expression vector by site-directed mutagenesis to obtain the double-mutant proteins F29M/F40M, F29M/L89M and F40M/L89M. The mutant proteins were expressed and purified successfully to typical yields of 2 mg (F29M/F40M), 1.2 mg (F29M/L89M) and 6 mg (F40M/L89M) per litre of medium. All three mutants showed oxygenase activity that was comparable to that of the native protein. Selenomethione-substituted samples of all three double mutants could be prepared; however, the final yields of soluble protein from the F29M/F40M mutant (0.6 mg l^{-1}) and the F29M/L89M mutant (0.3 mg l^{-1}) were significantly lower than that of the nonsubstituted protein. In addition, after the purification protocol the purity of these two mutants was compromised by a contaminating

Table 1
Crystallographic data for recombinant SnoaB.

Values in parentheses are for the highest resolution shell.

Space group	$P2_12_12$
Unit-cell parameters (Å)	
a	58.8
b	114.1
c	49.5
Resolution (Å)	57–2.4 (2.53–2.4)
Beamline	ID23-2
Wavelength (Å)	0.8726
No. of observations	52927 (6617)
No. of unique reflections	13482 (1829)
$R_{\rm merge}$† (%)	6.2 (53.4)
Completeness (%)	99.2 (95.5)
$\langle I/\sigma(I)\rangle$	15.2 (2.3)
Multiplicity	3.9 (3.6)
Wilson *B* factor (Å^2)	53.0

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the *i*th observation of reflection *hkl* and $\langle I(hkl)\rangle$ is the weighted average intensity for all observations *i* of reflection *hkl*.

protein of higher molecular weight. However, the selenomethionine-substituted F40M/L89M mutant enzyme could be prepared to homogeneity with SeMet substitution with a final yield of 15.7 mg (12 mg per litre of medium). Mass-spectrometric analysis confirmed the incorporation of two selenomethionines per monomer.

Recombinant SnoaB was crystallized in 26–28%(*w*/*v*) PEG MME 2K, 50 m*M* KBr with 2%(*v*/*v*) pentaerythritol ethoxylate (3/4 EO/OH) as an additive that was added either to the reservoir solution or directly to the drop. Crystals grew in a cluster of plates after approximately three weeks at 277 K (Fig. 3). Most crystals that were tested on a synchrotron-radiation source suffered from severe anisotropy in their diffraction pattern. However, some of the crystals were better ordered and a complete data set was collected from one crystal to a resolution of 2.4 Å (Table 1). Based on inspection of the diffraction data and on the auto-indexing routine of *MOSFLM*, space group $P2_12_12$ with unit-cell parameters $a = 58.8$, $b = 114.1$, $c = 49.5$ Å was assigned. A $V_{\rm M}$ of 1.98 Å^3 Da^{-1} would correspond to the presence of three molecules in the asymmetric unit with a solvent content of 37.8%. Alternatively, two molecules in the asymmetric unit would correspond to a solvent content of 58.5% and a $V_{\rm M}$ of 2.96 Å^3 Da^{-1} (Matthews, 1968).

The SeMet-substituted SnoaB double mutant F40M/L89M was crystallized under identical conditions as the wild-type protein. Crystals appeared within one week after streak-seeding from wild-type crystals. The crystals were further improved by secondary streak-seeding using the initial SeMet-substituted crystals as seeds. Preliminary data indicated that the selenomethionine-substituted F40M/L89M crystals belonged to the same space group as those of wild-type SnoaB and a more detailed analysis is in progress.

We gratefully acknowledge access to synchrotron radiation at MAX-lab, Lund University and the ESRF, Grenoble, France. This work was supported by grants from the Swedish Research Council and the Academy of Finland (SA 121688).

References

Adams, M. A. & Jia, Z. (2005). *J. Biol. Chem.* **280**, 8358–8363.
Beinker, P., Lohkamp, B., Peltonen, T., Niemi, J., Mäntsälä, P. & Schneider, G. (2006). *J. Mol. Biol.* **359**, 728–740.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Combet, C., Blanchet, C., Geourjon, C. & Deléage, G. (2000). *Trends Biol. Sci.* **25**, 147–150.
Cross, E. J. & Perkin, A. G. (1930). *J. Chem. Soc.*, pp. 292–308.
Cuff, J. A. & Barton, G. J. (2000). *Proteins*, **40**, 502–511.
Fetzner, S. (2002). *Appl. Microbiol. Biotechnol.* **60**, 243–257.
Fetzner, S. (2007). *Nature Chem. Biol.* **3**, 374–375.
Fischbach, M. A. & Walsh, C. T. (2006). *Chem. Rev.* **106**, 3468–3496.
Gassner, N. C. & Matthews, B. W. (1999). *Acta Cryst.* D**55**, 1967–1970.
Hertweck, C., Luzhetskyy, A., Rebet, Y. & Bechthold, A. (2007). *Nature Prod. Rep.* **24**, 162–190.
Higuchi, R., Krummel, B. & Saiki, R. K. (1988). *Nucleic Acids Res.* **16**, 7351–7367.
Jämsen, J., Tuominen, H., Salminen, A., Belogurov, G. A., Magretova, N. N., Baykov, A. A. & Lahti, R. (2007). *Biochem. J.* **408**, 327–333.
Kallio, P., Sultana, A., Niemi, J., Mäntsälä, P. & Schneider, G. (2006). *J. Mol. Biol.* **357**, 210–220.
Leahy, D. J., Erickson, H. P., Aukhil, I., Joshi, P. & Hendrickson, W. A. (1994). *Proteins*, **19**, 48–54.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Metsä-Ketelä, M., Niemi, J., Mäntsälä, P. & Schneider, G. (2008). *Top. Curr. Chem.* **282**, 101–140.
Minotti, G., Menna, P., Salvatorelli, E., Cairo, G. & Gianni, L. (2004). *Pharmacol. Rev.* **56**, 185–229.
Schneider, G. (2005). *Curr. Opin. Struct. Biol.* **15**, 629–636.
Sciara, G., Kendrew, S. G., Miele, A. E., Marsh, N. G., Federici, L., Malatesta, F., Schimperna, G., Savino, C. & Vallone, B. (2003). *EMBO J.* **22**, 205–215.
Stec, B. & Stieglitz, K. A. (2008). *Acta Cryst.* D**64**, 1000–1002.
Sultana, A., Alexeev, I., Kursula, I., Mäntsälä, P., Niemi, J. & Schneider, G. (2007). *Acta Cryst.* D**63**, 149–159.
Widboom, P. F., Fielding, E. N., Liu, Y. & Bruner, S. D. (2007). *Nature (London)*, **447**, 342–345.
Ylihonko, K., Hakala, J., Kunnari, T. & Mäntsälä, P. (1996). *Microbiology*, **142**, 1965–1972.
Ylihonko, K., Tuikkanen, J., Jussila, S., Cong, L. & Mäntsälä, P. (1996). *Mol. Gen. Genet.* **251**, 113–120.