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

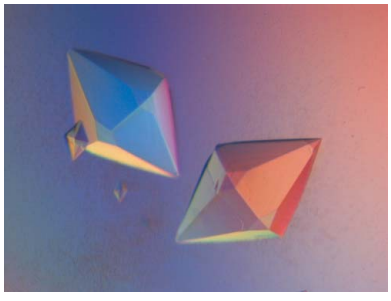
Crystallization and preliminary crystallographic analysis of mannosyl-3-phosphoglycerate synthase from *Rubrobacter xylanophilus*

Bebiana Sá-Moura,[a] Luciana Albuquerque,[b] Nuno Empadinhas,[b] Milton S. da Costa,[c] Pedro José Barbosa Pereira[a] and Sandra Macedo-Ribeiro[a]*

[a]IBMC – Instituto de Biologia Molecular e Celular, Universidade do Porto, Porto, Portugal, [b]Centro de Neurociências e Biologia Celular, Departamento de Zoologia, Universidade de Coimbra, Coimbra, Portugal, and [c]Departamento de Bioquímica, Universidade de Coimbra, Coimbra, Portugal

Correspondence e-mail: sribeiro@ibmc.up.pt

Received 11 May 2008
Accepted 10 July 2008



Rubrobacter xylanophilus is the only Gram-positive bacterium known to synthesize the compatible solute mannosylglycerate (MG), which is commonly found in hyperthermophilic archaea and some thermophilic bacteria. Unlike the salt-dependent pattern of accumulation observed in (hyper)thermophiles, in *R. xylanophilus* MG accumulates constitutively. The synthesis of MG in *R. xylanophilus* was tracked from GDP-mannose and 3-phosphoglycerate, but the genome sequence of the organism failed to reveal any of the genes known to be involved in this pathway. The native enzyme was purified and its N-terminal sequence was used to identify the corresponding gene (*mpgS*) in the genome of *R. xylanophilus*. The gene encodes a highly divergent mannosyl-3-phosphoglycerate synthase (MpgS) without relevant sequence homology to known mannosylphosphoglycerate synthases. In order to understand the specificity and enzymatic mechanism of this novel enzyme, it was expressed in *Escherichia coli*, purified and crystallized. The crystals thus obtained belonged to the hexagonal space group $P6_522$ and contained two protein molecules per asymmetric unit. The structure was solved by SIRAS using a mercury derivative.

1. Introduction

The radiation-resistant bacterium *Rubrobacter xylanophilus* is moderately halotolerant and has an optimum growth temperature of about 333 K (Carreto *et al.*, 1996). It accumulates trehalose and mannosylglycerate (MG) intracellularly under both normal and stress growth conditions (Empadinhas *et al.*, 2007). Two pathways for the synthesis of MG have so far been identified: one involves the direct conversion of GDP-mannose and D-glycerate into MG by mannosylglycerate synthase (MgS), while the other comprises the consecutive synthesis and dephosphorylation of mannosyl-3-phosphoglycerate (MPG) by MPG synthase (MpgS; EC 2.4.1.217) and MPG phosphatase (MpgP; EC 3.1.3.70) from GDP-mannose and 3-phosphoglycerate (3-PGA) (Empadinhas *et al.*, 2001; Martins *et al.*, 1999).

Glucosylglycerate (GG), a solute structurally related to MG, has been detected in bacteria and archaea in minor amounts and has been shown to serve in the osmotic adaptation of *Erwinia chrysanthemi* in nitrogen-depleted media (Goude *et al.*, 2004). Two pathways for the synthesis of GG have so far been elucidated: a single-step pathway and a two-step mechanism *via* glucosyl-3-phosphoglycerate (GPG), mechanistically similar to those for the synthesis of MG (Costa *et al.*, 2006; Fernandes *et al.*, 2007). GG has also been detected in the polar head of a glycolipid from a species of *Nocardia* and in methylglucose polysaccharides from *Mycobacterium* sp. (Kamisango *et al.*, 1987; Tuffal *et al.*, 1998). A highly divergent GPG synthase (GpgS) has been recently characterized in *M. bovis* and in *M. smegmatis*, where it has been proposed to synthesize GPG as a primer for the assembly of unique GG-containing methylglucose polysaccharides (Empadinhas *et al.*, 2008; Kamisango *et al.*, 1987; Tuffal *et al.*, 1998). The mycobacterial GpgS gene has close homologues in many *Actinobacteria* and, coincidently, the *R. xylanophilus* homologue is the *mpgS* gene that has now been identified. This ancient enzyme might represent a close relative of the common ancestor of strict MpgSs and GpgSs preceding the divergence of these two activities. The closest

R. xylanophilus MpgS homologue of known three-dimensional structure is the spore-coat-forming protein SpSA from *Bacillus subtilis* (11% identity), a monomeric glycosyl transferase belonging to GT family 2 and possessing a GT-A-like fold (Charnock & Davies, 1999). Structural characterization of this novel enzyme by X-ray crystallography should provide further insights towards the clarification of its specificity and constitute a framework for a detailed structural comparison with homofunctional GpgSs. Here, we report the expression, purification, crystallization and preliminary crystallographic analysis of the MpgS from *R. xylanophilus*.

2. Materials and methods

2.1. Purification of native MpgS

R. xylanophilus DSM 9941 cells (Deutsche Sammlung von Mikroorganismen und Zellkulturen, Braunschweig, Germany) were grown at 333 K in *Thermus* medium pH 7.0 (Williams & da Costa, 1992). Cells were harvested during the late exponential phase of growth, suspended in 20 m*M* Tris–HCl pH 7.6 containing 5 m*M* $MgCl_2$, 2 mg ml^{-1} DNAse I and a protease-inhibitor cocktail (Roche, Germany), disrupted with a French press cell and centrifuged to remove debris. The supernatant was dialyzed against 20 m*M* Tris–HCl pH 7.6 (buffer *A*) and loaded onto two sequential Q-Sepharose FF columns (Hi-Load 16/10) equilibrated with the same buffer. Elution was carried out with linear NaCl gradients (0.0–1.0 *M*) and the enzyme activity was determined as described previously (Empadinhas *et al.*, 2001, 2008). Active fractions were concentrated, dialyzed against buffer *A* and further purified by ion-exchange chromatography on a Mono-Q column using the same buffer system. The purest sample was blotted onto a PVDF membrane and subjected to N-terminal amino-acid sequencing (Microchemical Facility, Emory University School of Medicine, Georgia, USA). The N-terminal protein sequences obtained were used to identify the MpgS gene in the *R. xylanophilus* genome sequence (http://www.jgi.doe.gov).

2.1.1. Cloning and functional overexpression of *mpgS* in *Escherichia coli*. Specific primers (RuMG1, 5′-CATGAATTCATGAGTACTTATCATGAGAG-3′, and RuMG2, 5′-CATAAGCTTCTACCTTACCCTCAGCACCTC-3′) were designed and used to amplify the *mpgS* gene (GenBank accession No. EU847586) by PCR using 100 ng *R. xylanophilus* chromosomal DNA (Rainey *et al.*, 1996) as a template. The amplified fragment was cloned into the pGEM T-Easy vector (Promega) for sequencing purposes. The *mpgS* ORF was cloned between the *Eco*RI and *Hin*dIII sites of pET30a (Novagen) to allow the production of an N-terminal His-tagged protein. The expression construct encodes all 335 residues of mature MpgS, preceded by 52 residues forming the affinity tag (MHHHHHHSSGLVPRGSGMKETAAAKFERQHMDSPDLGTDDDDKAMADIGSEF). Expression in *E. coli* BL21 (DE3) cells was induced with IPTG (final concentration 1 m*M*) at mid-exponential growth (OD_{600} = 0.8) and allowed to proceed for 6–8 h at 303 K. The cells were harvested and the crude extract prepared as described in §2.1.

2.1.2. Purification of recombinant MpgS. The His-tagged MpgS was purified by immobilized metal-affinity chromatography on an Ni-Sepharose column (His-Prep FF 16/10) equilibrated with 20 m*M* sodium phosphate buffer pH 7.4, 0.5 *M* NaCl, 20 m*M* imidazole. Elution was carried out with a linear gradient of imidazole (0.05–0.5 *M*) and MpgS activity was detected as described previously (Empadinhas *et al.*, 2001, 2008). The selected fractions were dialyzed against 20 m*M* Tris–HCl pH 8.0 and further purified by anion-exchange chromatograpy (Q-Sepharose FF). The purity of the samples was determined by SDS–PAGE.

2.2. Crystallization, data collection and processing

2.2.1. MpgS crystallization. Initial crystallization conditions were identified in sitting-drop trials using commercial sparse-matrix crystallization screens (Hampton Research). After optimization, the best native MpgS crystals were obtained at 293 K using the hanging-drop vapour-diffusion method from drops consisting of 4 µl protein solution (16 mg ml^{-1} in 10 m*M* Tris–HCl pH 7.5, 1 m*M* $MgCl_2$) and 2 µl precipitant solution (0.1 *M* MES pH 6.0, 2.0 *M* NaCl, 0.1 *M* NaH_2PO_4 and 0.1 *M* KH_2PO_4) equilibrated against a 500 µl reservoir. Crystals of the MpgS–GDP–mannose ternary complex were obtained from a 1:10:10 protein:GDP:mannose mixture pre-incubated at 277 K for 1 h using the crystallization conditions for the native protein. To obtain the mercury derivative, crystals of native MpgS were soaked for 12 h in a solution of *p*-chloromercuribenzoic acid (0.1 *M* MES buffer pH 6.0, 2.0 *M* NaCl, 0.1 *M* NaH_2PO_4, 0.1 *M* KH_2PO_4, 10 m*M* *p*-chloromercuribenzoic acid). The cryoprotection protocol involved the sequential transfer of crystals into drops of 0.15 *M* MES pH 6.0, 2.4 *M* NaCl, 0.15 *M* NaH_2PO_4 and 0.15 *M* KH_2PO_4 supplemented with increasing concentrations of glycerol [15, 20 and 25%(*v*/*v*)]. Crystals were finally cryocooled and stored in liquid nitrogen prior to X-ray data collection.

2.2.2. Data collection and processing. X-ray diffraction data for native MpgS and mercury-labelled MpgS were collected on beamline ID14-EH3 of the European Synchrotron Radiation Facility (ESRF, Grenoble, France). For the native protein, a data set extending to 2.20 Å resolution was collected on an ADSC Q4R detector from a single crystal at 100 K in 0.5° oscillation steps over a range of 150°. For the mercury derivative, a data set extending to 3.05 Å resolution was collected using a single crystal in 1° oscillation steps over a range of 120°. For the ternary complex with GDP and mannose, a data set extending to 2.80 Å resolution was collected at 100 K from a single crystal in 0.5° oscillation steps over a range of 135° using an ADSC 210 detector on the macromolecular crystallography beamline ID14-EH1 of the ESRF (Grenoble, France). Diffraction data sets were processed with *MOSFLM* (Leslie, 1991) and *SCALA* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994).

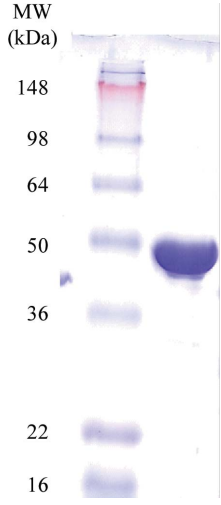


Figure 1
Coomassie Blue-stained 12.5% SDS–PAGE of the purified recombinant MpgS from *R. xylanophilus* used for crystallization trials.

Table 1
Statistics of diffraction data collection.

Values in parentheses are for the outermost shell.

	Native	Hg derivative	GDP–mannose complex
X-ray source	ESRF ID14-EH3	ESRF ID14-EH3	ESRF ID14-EH1
Wavelength (Å)	0.931	0.931	0.934
Temperature (K)	100.0	100.0	100.0
Space group	$P6_522$	$P6_522$	$P6_522$
Unit-cell parameters (Å)	$a = b = 108.9$, $c = 311.7$	$a = b = 108.7$, $c = 311.7$	$a = b = 109.0$, $c = 313.4$
Solvent content (%)	60.4	60.4	60.4
Resolution range (Å)	77.8–2.20 (2.32–2.20)	80.6–3.05 (3.21–3.05)	90.5–2.80 (2.95–2.80)
No. of reflections (total/unique)	1723540/56613 (147932/8094)	287757/21781 (41572/3088)	385577/27781 (56168/3948)
Multiplicity	30.4 (18.3)	13.2 (13.5)	13.9 (14.2)
Completeness (%)	100 (100)	100 (100)	99.7 (99.0)
$R_{\rm merge}$†	0.068 (0.604)	0.095 (0.417)	0.061 (0.311)
$R_{\rm r.i.m.}$‡	0.069 (0.634)	0.104 (0.446)	0.065 (0.334)
$R_{\rm p.i.m.}$§	0.012 (0.147)	0.028 (0.121)	0.017 (0.087)
$I/\sigma(I)$	7.9 (1.2)	7.1 (1.8)	10.7 (2.4)
Mean(I)/sd(I)	34.9 (5.2)	26.3 (5.6)	28.7 (10.7)
Anomalous completeness (%)		100 (100)	
Anomalous multiplicity		7.2 (7.1)	
No. of derivative sites		2	
Phasing power (iso/anom)		0.997/0.904	
$R_{\rm Cullis}$ (iso/anom)		0.836/0.854	
Figure of merit (acentric/centric)		0.141/0.183	

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity and $\langle I(hkl)\rangle$ is the average intensity of multiple observations of symmetry-related reflections. ‡ $R_{\rm r.i.m.} = \sum_{hkl}[N/(N-1)]^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity and $\langle I(hkl)\rangle$ is the average intensity of multiple observations of symmetry-related reflections. § $R_{\rm p.i.m.} = \sum_{hkl}[1/(N-1)]^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity and $\langle I(hkl)\rangle$ is the average intensity of multiple observations of symmetry-related reflections.

2.3. Structure solution

The structure was solved by the single isomorphous replacement with anomalous scattering (SIRAS) method. Using *autoSHARP* (Vonrhein *et al.*, 2007) and data from the native crystal and the mercury derivative, two mercury sites were successfully located. Phases were further improved and extended to the diffraction limit of the native data by density modification and twofold averaging using *SOLOMON* (Abrahams & Leslie, 1996) and then used for automatic model building in *ARP/wARP* (Perrakis *et al.*, 1999).

3. Results and discussion

3.1. Protein characterization and recombinant expression

Native MpgS was partially purified in four chromatographic steps. The purest preparation contained four bands on SDS–PAGE, the N-terminal sequences of which were determined (N. Empadinhas & M. S. da Costa, unpublished results). These sequences allowed the identification of four distinct ORFs in the *R. xylanophilus* genome, one of which (Rxyl_2312) coded for a 37.3 kDa putative family 2 glycosyltransferase, a likely candidate for MpgS activity.

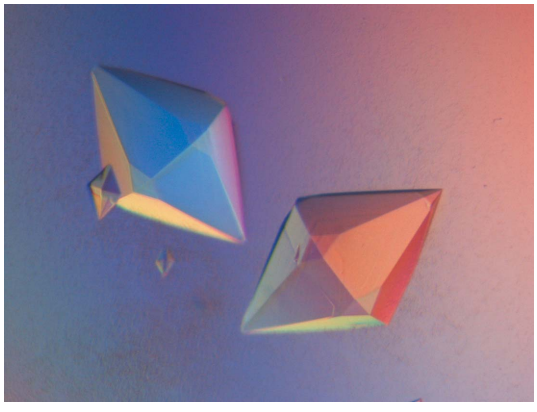


Figure 2
Single crystals of native MpgS from *R. xylanophilus* belonging to the hexagonal space group $P6_522$.

Cloning and expression of Rxyl_2312 in *E. coli* produced clones whose protein extracts showed extra bands of ∼43 kDa on SDS–PAGE (consistent with the molecular weight of the fusion construct) and catalyzed MPG synthesis, while extracts from *E. coli* with the empty vector did not. Heat-treatment (333 K, 20 min) of independent *E. coli* extracts containing His-tagged MpgS resulted in extensive purification of the recombinant enzyme and further purification by FPLC yielded pure protein as judged by SDS–PAGE (Fig. 1). The recombinant His-tagged protein eluted from the gel-filtration column with an apparent molecular weight of approximately 95 kDa, suggesting that it is dimeric in solution, which was confirmed by DLS (data not shown). Biochemical assays to determine the enzymatic activity and specificity of MpgS are currently under way.

3.2. Crystallization of MpgS

Crystals of native MpgS were hexagonal bipyramidal with maximum dimensions of 0.6 × 0.25 × 0.3 mm (Fig. 2) and diffracted to 2.2 Å resolution (Fig. 3). The crystals were extremely fragile and manipulation for cryoprotection often resulted in reduced diffraction quality. In order to obtain improved diffraction patterns, quick transfer (less than 5 s) between the different cryoprotecting solutions containing increasing concentrations of glycerol was critical.

The data-collection and phasing statistics are summarized in Table 1. The crystals belong to the hexagonal space group $P6_522$, with unit-cell parameters $a = b = 108.9$, $c = 311.7$ Å. Assuming the presence of two monomers per asymmetric unit, the calculated Matthews coefficient is 3.10 Å^3 Da^{-1}, which corresponds to a solvent content of 60.4% (Matthews, 1968). This relatively high solvent content explains the inherent mechanical fragility of the MpgS crystals.

3.3. Structure solution

The structure was solved by the single isomorphous replacement with anomalous scattering (SIRAS) method using *autoSHARP* (Vonrhein *et al.*, 2007). The hexagonal space-group ambiguity was

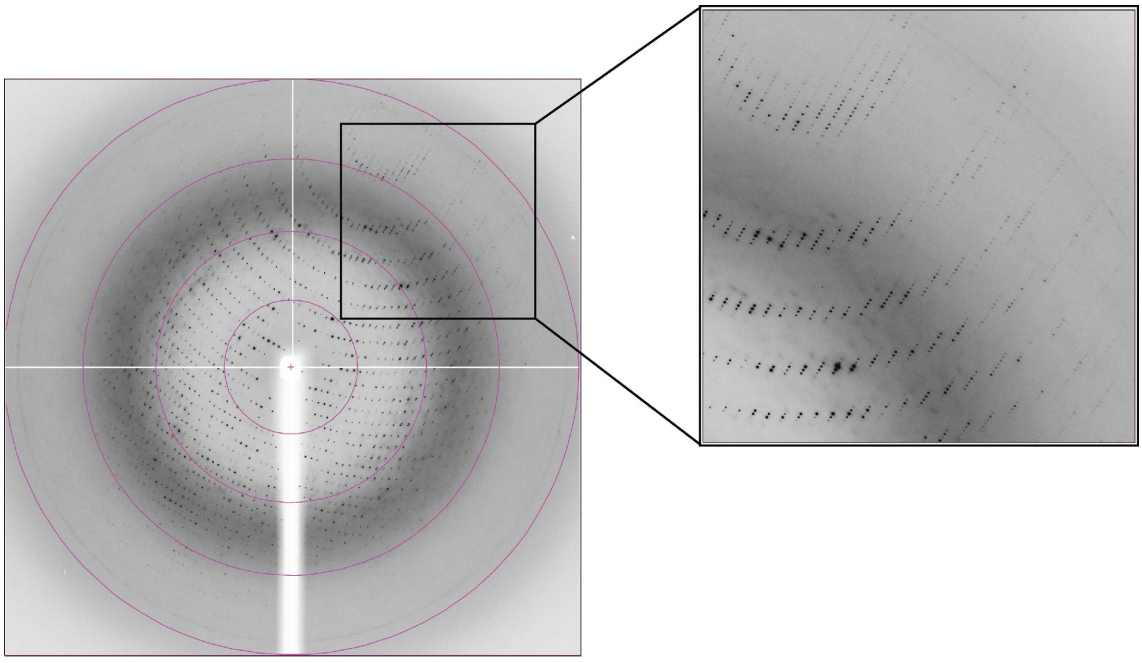


Figure 3
X-ray diffraction pattern of an MpgS crystal recorded on the ESRF ID14-EH3 beamline. The external circumference corresponds to 2.09 Å resolution and the resolution at the corner of the image is 1.59 Å. A magnification of the upper right quadrant of the image is shown to the right.

solved by analyzing the 'map contrast' and 'map connectivity' statistics calculated for space groups $P6_522$ and $P6_122$. The statistics and final electron-density maps were clearly better for $P6_522$ than for the alternative $P6_122$ space group. Given the relatively high solvent content of the crystals (60.4%), solvent flattening was highly effective at improving the quality of the maps, which was further helped by twofold averaging and phase extension to the resolution limit of the native data. The experimentally phased, solvent-flattened and twofold-averaged electron-density maps were of good quality and a significant portion of the model (583 of 774 residues) could be traced automatically with *ARP*/*wARP* (Perrakis *et al.*, 1999). The model is currently being refined and revealed a dimeric arrangement of the enzyme. The three-dimensional structure of *R. xylanophilus* MpgS will provide the basis for further studies aiming at understanding the structure and activity of this novel enzyme.

We acknowledge the ESRF for provision of synchrotron-radiation facilities and thank the ESRF staff for help with data collection at beamlines ID14-EH3 and ID14-EH1. We gratefully acknowledge funding from Fundação para a Ciência e Tecnologia, Portugal (POCI/BIA-MIC/56511/2004 and PTDC/SAU-MII/70634/2006).

References

Abrahams, J. P. & Leslie, A. G. W. (1996). *Acta Cryst.* D**52**, 30–42.
Carreto, L., Wait, R., Nobre, M. F. & da Costa, M. S. (1996). *J. Bacteriol.* **178**, 6479–6486.
Charnock, S. J. & Davies, G. J. (1999). *Biochemistry*, **38**, 6380–6385.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Costa, J., Empadinhas, N., Goncalves, L., Lamosa, P., Santos, H. & da Costa, M. S. (2006). *J. Bacteriol.* **188**, 1022–1030.
Empadinhas, N., Albuquerque, L., Mendes, V., Macedo-Ribeiro, S. & da Costa, M. S. (2008). *FEMS Microbiol. Lett.* **280**, 195–202.
Empadinhas, N., Marugg, J. D., Borges, N., Santos, H. & da Costa, M. S. (2001). *J. Biol. Chem.* **276**, 43580–43588.
Empadinhas, N., Mendes, V., Simoes, C., Santos, M. S., Mingote, A., Lamosa, P., Santos, H. & Costa, M. S. (2007). *Extremophiles*, **11**, 667–673.
Fernandes, C., Empadinhas, N. & da Costa, M. S. (2007). *J. Bacteriol.* **189**, 4014–4019.
Goude, R., Renaud, S., Bonnassie, S., Bernard, T. & Blanco, C. (2004). *Appl. Environ. Microbiol.* **70**, 6535–6541.
Kamisango, K., Dell, A. & Ballou, C. E. (1987). *J. Biol. Chem.* **262**, 4580–4586.
Leslie, A. (1991). *Crystallographic Computing 5*, edited by D. Moras, A. D. Podjarny & J. C. Thierry, pp. 27–38. Oxford University Press.
Martins, L. O., Empadinhas, N., Marugg, J. D., Miguel, C., Ferreira, C., da Costa, M. S. & Santos, H. (1999). *J. Biol. Chem.* **274**, 35407–35414.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.
Rainey, F. A., Ward-Rainey, N., Kroppenstedt, R. M. & Stackebrandt, E. (1996). *Int. J. Syst. Bacteriol.* **46**, 1088–1092.
Tuffal, G., Albigot, R., Riviere, M. & Puzo, G. (1998). *Glycobiology*, **8**, 675–684.
Vonrhein, C., Blanc, E., Roversi, P. & Bricogne, G. (2007). *Methods Mol. Biol.* **364**, 215–230.
Williams, R. A. D. & da Costa, M. S. (1992). *The Prokaryotes*, edited by A. Balows, H. G. Trüper, M. Dworkin, W. Harder & K. H. Schleifer, pp. 3745–3753. New York: Springer.