Molecular characterization of factors involved in regulation of archaeal translation

Fabian Blombach

Thesis committee

Thesis supervisor
Prof. dr. John van der Oost
Personal Chair at the Laboratory of Microbiology
Wageningen University

Thesis co-supervisor
Dr. ir. Stan J.J. Brouns
Researcher at the Laboratory of Microbiology

Other members
Prof. dr. Sacco C. de Vries, Wageningen University
Prof. dr. Bettina Siebers, University Duisburg-Essen, Germany
Dr. Finn Werner, University College London, United Kingdom
Dr. Berend Snel, Utrecht University

This research was conducted under the auspices of the Graduate School VLAG

Molecular characterization of factors involved in regulation of archaeal translation

Fabian Blombach

Thesis
submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. dr. M.J. Kropff,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Wednesday 19 May 2010
at 4 p.m. in the Aula.

Fabian Blombach
Molecular characterization of factors involved in regulation of archaeal translation, 142 pages.

Thesis, Wageningen University, Wageningen, NL (2010)
With references, with summaries in Dutch and English

ISBN 978-90-8585-643-6

Preface and outline of this thesis

This thesis deals with a number of proteins from Archaea with unknown roles in translation – the process in which the genetic information from nucleic acids is converted into proteins. Archaea form the third domain of life next to Bacteria and Eukaryotes. Many archaeal proteins that are part of the translation apparatus show a high degree of conservation in all cellular life. However, some translation-related proteins from Archaea only have counterparts in Eukaryotes, not in Bacteria. For the proteins under investigation in this thesis a function in the translation apparatus could only be inferred from published data about their bacterial or eukaryotic orthologs, but the high conservation of these proteins suggests that they play key roles in the assembly and regulation of the translation machinery. Two of these proteins - an HflX-type GTPase and MBF1 - have been characterized biochemically in this thesis. For a third one an in silico analysis has revealed that it is the archaeal ortholog of a eukaryotic protein involved in the transcription of structural RNA components of the translation apparatus.

Translation factor related GTPases and the HflX family

Chapter 1 reviews the knowledge about GTPases from the translation factor-related class in Archaea. These proteins can be predicted to fulfill various functions within the translation machinery. In all cellular life, a set of strictly conserved GTPases functions as factors that regulate the initiation and elongation steps of translation. In addition, other GTPases of this class function in the assembly of the ribosome, the large molecular machine that is the central component of the translation apparatus that catalyzes protein synthesis. **Chapter 2** describes the structural characterization of an archaeal HflX GTPase that belongs to the translation factor related GTPases. The structure reveals how this GTPase could possibly interact with the ribosome and highlights conserved features of HflX GTPases that can also be found in Eukaryotes and Bacteria. In **Chapter 3** the interaction of the archaeal HflX GTPase with the large ribosomal subunit is studied using biochemical methods.

MBF1

In Crenarchaeota the HflX GTPase is co-expressed with the helix-turn-helix protein MBF1. MBF1 is highly conserved in Archaea and Eukaryotes, but

absent in Bacteria. In yeast, gene deletion of *MBF1* leads to an increased rate of errors during translation, indicating that the protein influences directly (by binding to the ribosomes during translation) or indirectly (e.g. through a role in ribosome assembly) translation fidelity. The archaeal MBF1 ortholog remained uncharacterized so far. **Chapter 4** reports on the effects of *mbf1* gene deletion in Archaea. The *mbf1* deletion strain is less robust under a number of growth conditions when compared to the parental strain. In **Chapter 5** the MBF1 ortholog is shown to be a translational protein that binds directly to the ribosome during translation.

An archaeal RPC34 ortholog

In **Chapter 6** an archaeal gene is described that was found to be an ortholog of eukaryotic RNA polymerase III subunit RPC34. RNA polymerase III is one of the two RNA polymerases responsible for transcription of the structural RNA components of the translation machinery, ribosomal RNA and transfer RNA. The transcription of structural RNA differs from the transcription of messenger RNA regarding the different post-transcriptional processes that in turn can carry out feedback control functions for the transcription process. Little is known how Archaea organize the transcription of structural RNA, and it is not clear if there are differences between the transcription of messenger RNA and structural RNA. The archaeal RPC34 ortholog might be a factor that is specifically involved in the transcription of structural RNA, as its eukaryotic ortholog.

Chapter 7 presents a summary of the results described in this thesis, with some suggestions for future research.

Table of contents

CHAPTER 1 (INTRODUCTION)

Assembling the archaeal ribosome: potential roles for translation factor-related GTPases

manuscript in preparation

??rat? a

?A???aa?l ?lf?n??hs?nanl ?b?ac?ci sra??hnl ?rA?sh?si ?sos?c?b??nl gni?i ra?N?????
?i ??hs?nanl ?b?ghnr?si a)?h?t csh?a?rA???aasar?i ???h????l cbrsrc????h?????rmha?si ?
nh??h?n??ni rhnb??i ??si ?h??a??rA????s?s?i ?f?n??rA???aa?l ?lf?ghn??aa?P?????a?a?
n??rA??rh?i ab?rsni ????rmh?h?b?r????b?aa?N???????)??ni arsrcr????l ?enh?rfg??n??
hs?nanl ???aa?l ?lf????rmha?si ??cm?hfnr?a??i ????r?hs?P??A?f?h???tAnc Ar?n??s??
rA??ar?g: sa???aa?l ?lf?n??hs?nanl ?b?ac?ci sra?rAhnc A????vl nb???cb?h?a: sr?A(?
l ??A?isal ?rA?r?sionlo?a???ni?hhl ?rsni?b??A?i ?a?si?h?agni a???rn?????
Af?hnlfasaP??nar???ni a?ho?????????????????a?a???h???sionlo????si?hs?nanl ??
?aa?l ?lf?n?h?nrA?h?rh?i ab?rsniyh?b?r???ghn??aa?aP??A?f?rfgs??bf?si r?h???r? sr?A?
hs?nanl ?b?ac?ci sra???i??si?anl ????a?a?rA??si r?h???rsni ?A?a?????i?arc?s?????r?
l nb???cb?h?b?o?b??n: ?o?h?si ?l ?i f???a?a?rA???p??r?hnb??rA?r?rA?a?????a?a?gb?f?
h?l ?si a?ci ?b???h??h?osncalf?h?gnhr???arc?s?a??b narp??p?hcasolf?h?ca????hi?rA??
afar?l a?n?????r?hs???i ???cm?hfnr?al??h?A????gnaa?aa?a?o?h?h?ni a?ho???
??????????a?a??a? ?bf? srA?anl ?????a????l sbs?a???si ?gh?a?i r?nilf?si?rA??
?h?A???ny?cm?hfnrs??bsi ?? ?P????h??: ??h?os?: ?rA??n??chh?i ????n?????????
????a?a? srA?h?i ab?rsniyh?b?r????ci?rsni a?si??h?A????P

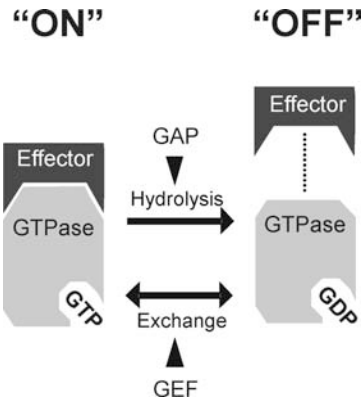
Introduction

Translation of the genetic information from nucleic acids into proteins is a highly conserved process in all three domains of life - Eukaryotes, Bacteria, and Archaea. The key player in this process is the ribosome, a large ribonucleoprotein complex composed of 3-4 distinct ribosomal RNAs (rRNAs) and 550-870 ribosomal proteins [1]. The ribosome catalyzes the formation of peptide bonds during protein synthesis, a process that requires the assistance of translation factors to achieve precision and efficiency during the initiation, elongation, and termination steps of translation. Some translation factors are universal, i.e. found in the three domains of life. Archaea employ translation factors that are a subset of the characterized eukaryotic translation factors whereas only the universally conserved factors are shared with Bacteria [2].

Several translation factors are GTPases, constituting the class of so called translation factor related GTPases (TRAFAC)[3]. This tight functional association of TRAFAC GTPases with the translation machinery suggests that they have co-evolved with the translation machinery. In fact TRAFAC GTPases are also involved in other translation-related processes such as tRNA modification and ribosome assembly [3]. Apart from the above mentioned translation factors, a few TRAFAC GTPase families are considered to be universal, while several other GTPase families are widely distributed in two of the three domains [4]. Little is known about the function of most of these highly conserved GTPases [5]. Several proposals have been made concerning the classification of the TRAFAC GTPase class, [3, 4, 6].

G-domain

The common denominator of GTPases is the highly conserved G-domain responsible for binding and hydrolysis of guanine nucleotides. Structurally, the G-domain of TRAFAC GTPases is characterized by a central β-sheet with a single β-strand in anti-parallel orientation. Conserved sequence motifs for the binding of guanine nucleotides and magnesium ions mainly reside on the loops and α-helices connecting the β-strand [3, 7, 8]. GTPases are considered to work as “molecular switches”: they undergo conformational changes when switching between the GTP- and the GDP-bound form, corresponding to the “ON”- and “OFF”-state of the GTPase, respectively [9]. Several factors take part



TAF f Og R ") A u t rG a Rf) R 1 0) A r r rw m C1 T aG 1 aR f aR f fra a A A1 atA C aR N 1 Cam A aR A t (aAt) A u C w r r rc uC a Tu 1 C 1u C Ac Rn tA Cm O R T et rb ct C v

si rA h cb rsni h rAsa a: sr Asi A i sal Ns Pk) P A v (y i v (y ar r h r A a h i a c si rn s h i r arhc rc h b ni h hl rsni a h rA r: n l n sh h sni a h A y nl si G h l a: sr A P rmha h si a rmha si si rn rA a sr As A h si sf si rA v (yar r nl g h rn rA v (yar r qk Ef P a rso rsi ghnr si a N a) arsl cb r Af hnbfasa f tA a a P c i si t c b nrs p A i rmha N a) ghnl nr rA p A i n c i si i c b nrs a nci rn rA a P m cm hf nr aC al b yghnr si a si onlo si as i b rh i a c rsni ac A a gBk a h h cb r f a i a si arhsr l i i h qkkf C cr rA a a si onlo si hs nanl aa l bf aAn: ni as h bf bn: h si sf hh c i si i c b nrs a a p l gbs3s f rA ar p A i msi rs a h rA wE hs nanl ba ac c i sr aa l bf a h h hnl qk Bf A h h h C sr a l a ci bsm f rA r A a a h g i i r hi a a s qkwf rA a a h hhs nanl aa l bf a a C hs nanl ba ac c i sra A o i aAn: i rn gb f h nba a rmha qk" f h a a qk" k M f A f banf s Ar m m o h nrA t i rsni a asl cb r i ncab f qk Hykx f l

A l enhsrf n hs nanl y aan s r a a h l cbrsy nl si ghnr si a si As A rA srsni b nl si a h si onlo si s i si r h rsni a l chsncab f C brAnc A hs nanl b ac c i sra h rA nl l ni s i h rA a a a a C rA srsni b nl si a rA r l s r rA si r h rsni h i ni y nhrAnbm nca l a nl a a A a A a h b bymin: i y si si nl si a ac A a y nl si a N l K qBEf) C A h a si nrA h a a rA f h h c i st c nl si a : Ana n chh i sa h arhsr rn rAsa a l sf N l K B i C rAsa A asa) P

:

Ribosome assembly in Archaea: prokaryotic and eukaryotic features

Ribosomes are ribozymes: it is essentially the rRNA that catalyzes peptidyl transfer. In addition, rRNA provides major elements of the kinetic proof reading mechanism [21]. The positively charged ribosomal proteins neutralize the negative charges of the phosphate backbone of rRNA, and as such contribute to compaction of the ribosome structure. Beyond their structural role, ribosomal proteins also have additional functions for example in tRNA and mRNA binding and translation factor recruitment [22].

The high number of ribosomal proteins that need to be incorporated into the ribosome requires the assembly process to be organized in a stepwise manner. Both in Bacteria and Eukaryotes, GTPases function in the organization of a stepwise assembly by serving as checkpoints that regulate the recruitment of additional ribosome assembly factors or ribosomal proteins to ribosome precursors [13]. Ribosome assembly is greatly facilitated by usage of the "assembly gradient", at least in case of the large ribosomal subunit. In this process, the first ribosomal proteins bind already during transcription to the 5'-end of the growing precursor rRNA transcript that later becomes processed to the 23S rRNA in Prokaryotes and the 28S rRNA in Eukaryotes [23]. Thereby, a 5' to 3' order is established for the incorporation of ribosomal proteins into the large ribosomal subunit.

In Eukaryotes the nucleolus is the location where ribosome biogenesis starts with the transcription of rRNA genes by RNA polymerases I (18S, 5.8S and 28S rRNA) and III (5S rRNA) that are dedicated to the transcription of non-coding RNAs. The evolution of dedicated RNA polymerases for non-coding RNAs might have its origin in the "archaeal parent" of Eukaryotes, as an archaeal ortholog of the RNA polymerase III subunit RPC34 that was assumed to be unique for this RNA polymerase, has been recently identified in Cren- and Thaumarchaeota [24] **(Chapter 6)**. Genome context analysis further suggests that the archaeal RPC34 ortholog might be involved in the transcription of non-coding RNAs. The compartmentalization of the eukaryotic cell with separate locations for the transcription of rRNA and the translation of ribosomal proteins poses a problem for ribosome assembly. Ribosomal proteins have to be imported to the nucleolus in order to make use of the "assembly gradient". Subsequently, the ribosomal subunit precursors

aal lhsi rA ic bnbca Aompgnhrsi rnrA frngbal Ghgnaa rAr ghnaosr: nsshi rhncrahhrA hi ri rA al bhhsnanl hb ac ci srP

nhrA i BEE ghnrsi asirs3srmrrhsionbosi cmhfnrs hsnanl aal lhf pB" GBMfPi lf al bhl si nhsrf hrAnarmhaAn: ahhhh nhrAnbn f rnhA bi aGsi hc si rA y cs l Asi hsahhhf si ncrrA l ns3srsni ahhhhi cbhnrsaapBHfPhA i gh srrrn l gbnf : hhrmhahhhsnanl aal lhf nl g hhrmcmhfnrahrr sbh sir hrarsi rnsio ars rrrm: Arrpri rhA bhsnanl aal lhf h al bhaA nhhagni si ghnaasi rrrhsqB, fP

l wsarA ni bf hsnanl pgnhrrmhhhnl cmhfnrahhhA hnhs si i sai onbosi pgnhrhrA hh hsnanl bac ci srghchanhahhnl rA i cbhcapBxfPl wnhrAnbn ahhhmci si hi yi chfhA nr Gcr hrrai r si A cl y i nhhA nr Pl srA n hhhhpgnhrhhhsnanl bh ac ci sra si hA Gsrsaabsmbf rA r l wA a i srsni bhti rsni si rA aal lhf nhrA bh hsnanl bac ci srPi r hrarsi bf GrA hA bl wnhrAnbn sa ta rmnl si hbhr rmy" qwEf: As A sani nhrA ci so ha bhi ab rsni bmi rsni rmha qwkfP

R r(? rOk? at r(A1 aAtr ((?

A a l sf ah bhaas bh rhi ab rsni rmha ni r si arA mch ci so ha h ni a ho rh i ab rsni rmh ac l sbs a srA rA sh hA bh gh a i r rso a ykα GBR yBγR h(i BR" N bk)N s PB) A sh ti rsni a si rh i ab rsni si srs rsni i bmi rsni h hh bh rso bf bh A h r hs- P

γ sa ac ci sr h rA si srs rsni rmh R B h h A i cmhfnraGcrsrsa nrgh a i r si r hs Pca R By γ sa bma bf h bh r rmb N s PB) G a a a hs bm: GrA R R By R b ac l sbf sa ni as h rm agh ci so ha bf qwfP i srs rsni rmh R B hhl a A r hnrhsl h srA sra yac ci sr i sr sa hn r i rh bf r: i rA α y i βyac ci sral PA a: sr A h sni h R By Ghi h hA r: n l n sh h sni a h y nl si a gb f a gsonr h hnb si rA nl gb p ni rhnbsi rA gnasrsni h rA By ac ci sr sbc arh rsi An: rA i cbhnrs y nci ar r n rA y nl si rh i a c a si rn arhc rch b ni hhl rsni b A i a qwBfP ra cmhfnrs nhrAnbn By γ sa h cbrr f n: nhrA hsi srs rsni rmha " i Bε rA r r a i (h ag rso bf qwwf(cr rA a gh nr si a h l saasi si

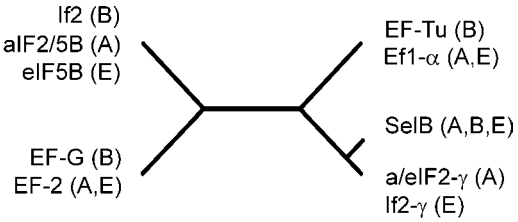


Figure 2: Schematic phylogenetic relationship between the universal GTPases from the translation factor family. Archaeal, bacterial, and eukaryotic occurrence is indicated by A,B, and E, respectively.

Archaea [2]. In its GTP-bound state, a/eIF2-γ recruits the initiator-tRNA to the 30S ribosomal subunit during translation initiation [34, 35]. *Sulfolobus solfataricus* a/eIF2-γ might have an additional function in RNA stability. It binds to the triphosphorylated 5' end of mRNA and protects it against 5'→3' degradation [36].

aIF2/5B is the archaeal ortholog of the eukaryotic translation factor eIF5B and bacterial IF2. In Bacteria, IF2 binds the initiator-tRNA to the 30S similar to the archaeal and eukaryotic heterotrimeric a/eIF2 described above. The archaeal IF2 ortholog aIF2/5B
stimulates the binding of initiator-tRNA to the 30S ribosomal subunit as well, but at least in *S. solfataricus* this does not involve a direct interaction with the initiator-tRNA [15]. The Ribosome might play a GAP-like function for the GTPase activity of *S. solfataricus* aIF2/5B [15].

Translation elongation factor **EF1-**α from Archaea and Eukaryotes binds as its bacterial ortholog EF-Tu aminoacyl-tRNAs in its GTP-bound state and delivers them to the ribosome during translation elongation. Recognition of proper base-pairing between the mRNA codon and the tRNA anticodon by the ribosome triggers GTP hydrolysis and release of EF1-α/EF-Tu. In all cellular life, EF-1α/EF-Tu depends on the GEF EF-1β/EF-Ts. However, the structure *S. solfataricus* EF-1α in a GDP-bound magnesium-free form pointed to possible differences in the mechanism of EF1-β- assisted guanine nucleotide exchange between Archaea and Eukarotes [37]. Data on the function of archaeal EF-1α in translation elongation are limited but overall confirm functional conservation across the domains of life. Interestingly, Archaea might use substrate channeling during translation elongation as mammalia do. *Methanothermobacter thermoautotrophicus* EF1-α forms complexes with aminoacyl-tRNA synthetases and in such way the unstable aminoacyl-tRNA synthesized by the aminoacyl-tRNA synthetase can be delivered directly to the ribosome [38].

A1

	P F1 2 f P l a	f o f P l a	F f b f P l a		l f f P l a
	r E' E'	E' c E' c E'		r E' E'	E' c d E'
P rr af or c lo al f 1 Tb					
R ByF	l hCEHB	E" kB EEM'	kHEM	kBMk	E" xB hCkEBB
BR"	l hCEBBx	EBB, k" kw	kEx"	EBMB	EHM, hCkEBE
yk.	l hCkEw'	EBkM wEEE	k", "	EwB"	ExwH hCk" xM
yB	l hCEH",	EHB, EwwB	k, xx	kE",	k, x" hCEkkH
b	l hCExMk	E" HE y	y	EwB"	y y
r r e f a l Fo c lo o f T l r l 1 rr 1 b					
k" M'	y	E", k EH' x	EM' "	k" M'	y hCEBEk
ap	y	y y	EBEE	EwBE	kwBM y
k	y	EMB" wMBM	kwBE	k" E,	E" B, hCk" ",
T l r l 1 2 To oA r r d Fo Ro l i o Fo c lo					
	l hCEM, "	Bw, " kM' x	kMHB	kwBM	Bk" M hCEEEk
hBkE	l hCk" w'	EH" w wkM'	E" B"	kwwB	kwM' hCkwH,
3t	l hCkME"	EBMk EH, B	kkwB	kkBM	y y
yk	y	B" Bx kHM'	y	y	EH" " hCk" " x
a P f f P e fl a To r T aP el a oc fl r o f T l r l 1 rr 1 b					
l w	y	Bw, w kM' w	ExxE	kM' E	kxHB y
	l hCE" wE	EHwH BBEx	k" " B	E" xB	E" xk hCkwHB
M	l hCEBwH	Ew" k wExH	E" B,	EE",	BEM' hCkwB,

si h A i cm hf nr a i rA r h s b nhr An bn y r h f - rA r Ah yi c b nrs r h i ab n r s ni n r A h s n an l n i r A l c h s i r h i ab r s ni n b m i r s ni P A h s n an l g b f a y b s m c i r s ni s i r A y B R y y a f b a n c b h an l n i ar h r h h h A b y B q w x C " E f P s r r b s a m i n: i n c r A c i r s ni b r r s b a n y B r h f - h s n an l r h i ab n r s ni i An: s r n l g h a r n s r a c m hf n r s i r h s b nc i r hg h r al P y B h m l r A g af Ah n g As b s c hf h A n i E A a a h o a l n b g h n r si h h t A n b g r r s ni h r A sa h h i sal q" Ey" BfP

Archaeal orthologs of the universal selenocysteine-specific translation elongation factor **SelB** are widely distributed in Cren-, Thaum-, and Euryarchaeota, but not strictly conserved (Table 1). Both thaumarchaeal genomes encode SelB orthologs, but they are missing in *Korarchaeum cryptofilum.* It functions analogous to EF1-α/EF-Tu by delivering during translation elongation a selenocysteine-charged tRNA to internal UGA codons present in certain mRNAs. UGA normally serves as one of the three stop codons. SelB orthologs from the Methanococcales have been studied in more detail. Gene disruption confirmed their involvement in selenoprotein synthesis [43]. The structure of the SelB ortholog from *Methanococcus maripaludis* confirmed structural homology of domains I to III of its four domains with EF1α/EF-Tu and IF2/eIF5B [44]. Structural homology between SelB and IF2 is extended to domain IV that might be involved in the binding to the mRNA to receive the recoding signal [44]. Interestingly, domain IV is conserved only in few SelB orthologs from methanogenic Archaea, but it is generally absent in other archaeal SelB orthologs. In addition, there is no strict co-occurrence of other components such as the selenocysteine-specific tRNA with SelB orthologs implicating that there is some variation in the mechanism of selenocysteine incorporation in the different organsims or alternatively that some SelB might have a different function.

Predicted ribosome assembly GTPases

The YlqF-related GTPase family and its archaeal member MJ1464

Various members of the different subfamilies of YlqF-related GTPases (YRG) family have been shown to participate in bacterial and eukaryotic ribosome assembly [19, 45-50]. A remarkable feature of the YRG family is the circular permutation of the G-domain. Bacteria and Eukaryotes possess several paralogous members of the YRG family. The MJ1464 subfamily is found only in Archaea where they are present in all Crenarchaeota and *Korarchaeum*, in some Euryarchaeota and (at present) in none of the Thaumarchaeota (Table 1). MJ1464 is the only archaeal representative within the family of YlqF-related GTPases (YRG) [3, 51].

Bacillus subtilis YlqF/RbgA loads the ribosomal protein L16 onto a 45S precursor of the large ribosomal subunit [50]. Similarly, the cytoplasmic

a a kg hml f ar sah t csh hhrA hn si hrA hs nanl hbghnr si kE hni rnrA hh hs nanl hbac ci srha hbh hhrA hh hbh a hhl wq " M l r hs hbk M i cm hf nrs kE h nhrAnbn nca ghnr si ac arsi rA r t R i a k l s Ar aA h i onbcrsni hf hi a ho ci rsni h a a l A f l s Ar gb f asl sh h hnh b si A r i a i a t c bsrf A m h rA arhc rch qkwf r sbh si r h arsi rn a A rA h k " M' sa si onbo si h si h kE hni rnrA si h h A P

A h hA bh k " M' a a i nl g aa yr hl si b nl si rA r sa Anl nbn nca rn rA yr hl si b α yA bs b nl si a h r hs b t a a q' Bf i rA cm hf nrs : f g a a rn As A rA hs nanl aa l bf a c k bni a q' Hf l Asa ci hbsi a rA gnr i rs b hnb n k " M' a a si hs nanl si asa

TP : oA : rn

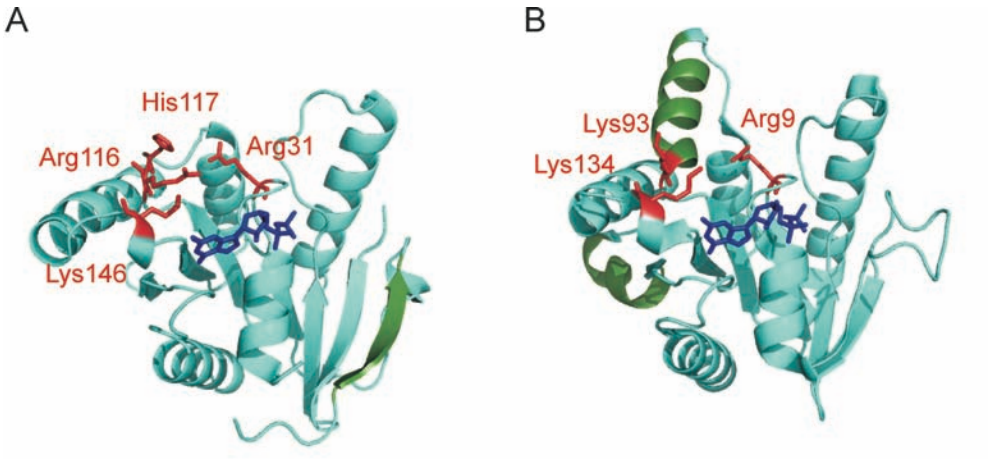
l ha h rA ap R i a l sf h c st csrnca si cm hf nr a i r hs C cr si h A rA sh n ch i sa h arhs r rn rA c hf h A h h i AC brAnc A rA f h ban l saasi si a o h b ag s a N b k) P A arhc rch h rA ap hrAnbn hml c E EBEE si rA y nci ar r A a i anbo f rA arhc rch b i nl s a si srs rso N cmsl nrny ssi n r r b C ci g c bsaA) N T B) P nl g hsani rn rA g c bsaA y nci ap arhc rch hml e E q' wf N s T k) h o ba ho h b nn arhc rch b ni a ho rsni Nhl a T k P) srAr: n srsni b α yA bs a si gh a i r si EBEE N s P w) P A 3shar A bsp sa si a hr r: i β" i α w Ne E ap ic l hsi n a ni hf arhc rch b l i ra) CrA a ni ni r: i α" i β HP Ash i ni s b y nl si a hml rA h b aa A o i rh b aspyarh i βyaA rC rA βyaA r h rA y nl si n ap ni r si a i srsni b yr hl si b βyarh i P A 3shar BH l si n s h asc a h BEE h l saasi si rA arhc rch C cr rA f hh hl i pr i asni rA r sa l nr ni a ho srAsi rA h A b ap hrAnbn a ac arsi rA r rA h A b ap hrAnbn a gnaa aa i ni s b aspyarh i βyaA rP gnasrso bf A h g r A ni rA ach h h ap ni asarsi h wk C kkM C kkH C i k " M q' wf a ghngna m si onbo si rA si r h rsni h ap srAr A b h hs nanl b ac ci srq' " f BEE rAsa g r A sa g hrs bf ni a ho N w" C kkx C i kME) As A l s Ar si s r ci rsni b ni a ho rsni P

NOG1

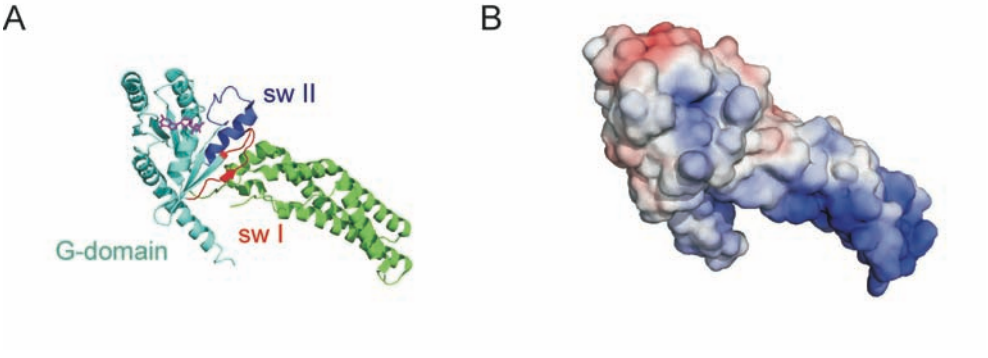
NOG1 GTPases are part of the Obg family of GTPases [3] and are highly conserved in Eukaryotes and Archaea except for Thaumarchaeota (Table 1). Eukaryotic NOG1 is a nucleolar GTPase involved in assembly of the large ribosomal subunit [46, 55]. Yeast NOG1 directly interacts with the ribosome assembly protein Rlp24, a eukaryotic paralog of the ribosomal protein L24e [56]. Archaea have only a single L24e ortholog that is part of mature large ribosomal subunits in *Haloarcula marismortui* [57]. Interestingly, the archaeal L24e ortholog shares features with Rlp24 such as two pairs of cysteines [56]. Thus, it seems possible that the archaeal L24e has a function in ribosome assembly as well including interaction with the archaeal NOG1 orhtolog. This would be similar to ribosomal protein L7ae that has a double life as structural component of the ribosome as well as ribosome assembly factor in the RNA-guided rRNA modification machinery [58].
The crystal structure of the NOG1 ortholog from *P. horikoshii* in GDP-bound form has been solved by the RIKEN structural genomics initiative, but it has not been published yet (PDP ID: 2E87). The N-terminal 160 amino acids comprise a four-helical bundle that provides positively charged surface patches that might be involved in ribosome binding (Fig. 4AB). The two mobile elements switch I and switch II of the G-domain mediate contacts with the N-terminal domain following a pattern commonly found in multi-domain GTPases for the interaction of the G-domain with a ligand binding domain [59-61]**(Chapter 2)** [20]. Hence, conformational changes of the switch regions in response to GTP-binding will most likely cause a repositioning of the N-terminal domain.

Ribosome-binding GTPases of unknown function

Apart from the classical translation factors and GTPases where a role in ribosome assembly has been demonstrated, several other TRAFAC GTPases interact with ribosomal subunits, but their biological role remains to be determined. For other TRAFAC GTPases no functional studies are available, but based on their domain composition a translation-related function can be predicted. Some of these poorly characterized GTPases are discussed below.



AFf q l A) b A1 aG 1 y g r- tA) l o i A1 At) y l, rtiG y gaR r- AtaRAC AT pxx tA) l o i B d A1 At) y lp- - gW4.R Cr h. rat r btr1aA1 mC A1 A aR aGA AtaRAC ATAur btAa C1 r AC At C1 t 1w R) AC uC C1 rRAG 1 C1 C u wR A1 r tti t rC u r At) C1 T bArC t C R T rut C1 l o i r- kvffi aR Att rbA1 C1 T t rC u r C1 aR l B d A1 AtaRAC AT RAG 1 C1 t w R R C1 r C1 C a t TC A1 r A aR btAa C1) C r C1 T C1 aR tmra ratu aut w



AFf E M atu aut A aR l B d A1 AtaRAC AT , vpx A aR ,) C y lp BSg y g GA. A) 1 tt T) 1a A AT, G aR .a t) C .A) 1 y n g .a t) C A) 1 At) C1 T Aut. R C u1 C y Tt 1g R aG A) A C) 1ar A aR .A) 1 aR rG aR t TC A1 re R RC Ra C1 t u e t rb ct C w y g R C at Ara rut t bt r1a A1 A1 aR rA ti 1a rr C rut A , vpx y., x U P aA , x U P g R I Tut G T 1 t a ur C1 n F6pff aR b U G a ur At u C A1 A at Ara rut bAa1c r F6pffiv

Obg family GTPases DRG and Ygr210

Next to Nog1, two more GTPases of the Obg family are present in Archaea: DRG and Ygr210 (Table 1) [3]. In both DRG and Ygr210 the N-terminal G-domain is followed by a C-terminal TGS domain [3]. This domain is also present in other Obg family GTPases and has been structurally characterized *B. subtilis* Obg [59] and *Haemophilus influenza* YchF [62]. Besides Obg family GTPases, TGS domains have also been found in Threonyl-tRNA synthetases and guanosine polyphosphatases (SpoT)[63].

Several bacterial Obg GTPases have been shown to associate with free 50S ribosomal subunits [64-66]. Obg GTPases control stringent response in Bacteria, but in addition they might play a role in ribosome assembly [64, 65, 67]. A thorough characterization of YchF is still missing, but a recent study of the *Trypanosoma cruzi* ortholog provides first evidence for interaction with the translation machinery [68]. The structure of *Haemophilus influenza* YchF further suggests binding of double-stranded nucleic acids [62]. Based on this circumstantial evidence it is likely that DRG and Ygr210 GTPases have translation-related functions as well. The G-domains of the Ygr210 family have canonical G4 sequence motifs conferring specificity for guanine nucleotides, in contrast to the closely related YchF subfamily in which mutated G4 motifs turn them into ATPases [69].

HflX

The HflX GTPase family is related to Obg GTPases [3]. HflX GTPases are ubiquitous in all three domains of life, but several taxonomic groups do not contain HflX orthologs. In Archaea, HflX GTPases are present in Crenarchaeota and in the available thaumarchaeal genomes, but absent from some euryarchaeal species (Table 1). The HflX family is the only TRAFAC GTPase family apart from the classical translation factors where experimental data about ribosome interaction is available from an archaeal representative. The structure of the HflX ortholog from *S. solfataricus* has been solved in apo- and GDP-bound form revealing the presence of a novel putative RNA-binding domain termed HflX-domain. This domain is sufficient to mediate the interaction with the large ribosomal subunit [70] **(Chapters 2 and 3)**. Similar to the bacterial HflX orthologs from *Escherichia coli* and *Chlamydophila pneumoniae* [71, 72] the archaeal HflX ortholog binds to the large ribosomal subunit both in the GTP- and GDP-bound form and its GTPase activity is

rso r cg ni si si P n si r h rsni : srA rh i ab rsi hs nanl a a h a ho h hrA h A b 3b h hrAnbn P si i h b rsni ah d s i nr h m i f h a ho g A i nrf g si r hs qHwyH" f sr h l si a ci h h A rA h d g b f a h mh b si hs nanl aa l bf P r hs rA g hnr si l s Ar nl sl gnhr i r i h arh aa ni srsni aqH" fP

i e E d sa nyrh i a hs : srA i qHMf i rA ni a ho rsni h rAsa i nl s ni r pr si b h i h A b i nl a ac ar rA r rA r: n i a h h f nyrh i a hs si h i h A nr qHHf P A h A b k h hrAnbn g g ha h m h i ab rsni b gh nr si sra b cr aan s r a ni bf i srA HE hs nanl a i h al b hs nanl b ac ci sr r h l si a ci h h A rA h rA h sa c i rsni b h b rsni aAsg r: i 3b i k si h A P

2Ot o Fo P f a f Tra r l aP r rT af or c l o al f

1 Tb

i nrA h a hnl rA b aas b rh i ab rsni rmh l sf rA r sa: s bf agh : srAsi h A sa yk qwf P r sa hc i si b h i h A nr a b a a o h b cm hf nrs i c hf h A b ag s aC cr sr sa a i r si A cl h A nr N b k)P o h b cm hf nr a gnaa aa l cbrsgb As Abf ni a ho yk g h bn aC A h a rA gh nr si sa a i r hnl anl cm hf nrs ag s alP n yk a nhrAnbn A a i rArAnhnc Abf A h r hs- C cr sr A a i ni arh r rA r nca yk bn r a m rA f rmgb al qH, fP yk a a nl ghsa i yr hl si b α y nl si rA r aAn: a n asl sh hsrf m hrA h ghnr si a r a tc i b o h r sa h bn: f i nis b y nl si i r: n nl si a Anl nbn nca rm nl si a i n k α R y c P A nl si nl gnasrsni b hf ac ara ci rsni si onbosi hs nanl si si a hmh nrA h a a hso hnl rA b aas b rh i ab rsni rmha ac A a h b nh g qHxfP

A cm hf nr yag s3s a 3kgR s kg sa h h r rmrh i ab rsni bni rsni rmh yB i gg ha m ci rsni m rA h srA rA A: Al i y n s i y s l ni f i hnl ghnr si N) casi arhc rc h b h hh i l i r si h r ar ghch anh m h rA h h hs nanl b ac ci sr b si rm rA h h b a h rA hs nanl aa l bf rmh s Mg q Ey, Bf P nrA i s Mg h hni a ho si cm hf nr a i h h A i sra b sa arhc rc h bf h b r rm nl si h yB q wf P A h A b s Mg h nhrAnbn M hnl E A a h i rf i aAn: i rn si rm rA b h hs nanl b ac ci sr

and to prevent subunit association [84] similar to its yeast ortholog [85]. Elf1 is absent in Archaea and the release of Tif6 from the large ribosomal subunit must thus be regulated in a different manner.

Conclusions

Archaeal genomes encode a number of TRAFAC GTPases with translation-related functions. The absence of NOG1, GP-1, and YRG family GTPases in Thaumarchaeota further corroborates the recent re-classification of these "marine Crenarchaeota" as a separate phylum next to Euryarchaeota and Crenarchaeota [86]. The number of TRAFAC GTPases encoded in thaumarchaeal genomes is surprisingly small with only three TRAFAC-GTPases (Ygr210, HflX, and DRG) serving as candidates for ribosome assembly GTPases.

The use of GTPases in ribosome assembly might allow the cells to gain control over the process of ribosome assembly [13]. Decreasing GTP levels during cellular stringent response to stress conditions would couple ribosome assembly directly to the energy state level of the cell. Interestingly, stringent response in *Sulfolobus* species appears not to correlate with decreasing GTP levels [87] suggesting that at least in the *Sulfolobus* genus ribosome assembly might not be regulated in this way.

Both in Eukaryotes and Bacteria, ribosome assembly GTPases predominantly act on the large ribosomal subunit [28]. The reason might be the considerably more complicated structure of the 23S rRNA. Similarly, two AAA ATPases (Rix7p and Rea1p) have been proposed to assist in large ribosomal subunit assembly in Eukaryotes [28]. So far no AAA ATPases have been identified that are involved in small ribosomal subunit assembly. Rix7p and Rea1p work as chaperones and a chaperone-like function has recently been proposed for the bacterial small subunit interacting GTPase Era [88], although direct evidence for such an activity has not been provided for any of the ribosome assembly GTPases [13].

Gene deletion or conditional gene knockdown in the case of essential genes have greatly contributed in the identification of ribosome assembly factors based on the accumulation of rRNA precursors as well as ribosomal subunit precursors that are unable to join translation. The conditional gene knockdown of ribosome assembly factors in combination with tap-tag

g c hs3s rsni n hhs nanl bg h c hanha A a bm: hh hr A r hl si rsni n hr A nl g nasrsni n rr Ana nrA h: sa bm: y c i i r g h c hanha q' M, x fl i r g hn h aa si rA 3s b n h A bo i rs a qx E(x k f : sb bm: hh rA s i rs3s rsni n h aa i rs bhs nanl aa l bf rmhal rA sa l i c a hsg r nl h A b hs nanl aa l bf rmha A o i g h s r a ni rA sh nhrAnbn f sr A c m hf nrs rmha N b k)l

A bc b h hnl g hrl i r bs- rsni n h c m hf nrs ba l hma aA hg s h i si hs nanl aa l bf r: i c m hf nr a i h A l rc s a n h rA h A b hnhrAnbn a n c m hf nrs hs nanl aa l bf rmha i n rA sh t i rsni si gh nm hf nrs hi r pr h bs m hf m n h o hl nh si as Ar si rn rA h l nh nl gb p c m hf nrs hs nanl aa l bf i si rn rA onbc rsni n nl g hrl i r bs- rsni n h A c m hf nrs bP

t 1 r

, w A) ba e Ob e RO ttn e At e A R l **l 1 e f c G o br T l f T l r l 1 efl a Tor o l 1 e a A ol 1 rg o n 1 e F c G Gl Fcl o a P l 1 To r u** / *d* pxxpe qwg kvBp.kv6xw

pw A1 Gl **Gl Fcl o af or c l o To T c clog o i TorTAPar fl 1 aP f P u** *u* pxxke YBg Bk.pxxw

vw Ob e AC AA1O e t tiO l r rrTC clo o oGl Fcl o 2 l le r r r o a a r r r u pxxpe qO6g , .Spw

: w C A1 e e R l Fo c l o aP FoTG fr b l or fG a f T r r r u / *BB* pxxve 4g vk., v6w

kw b tO e AA1O l l or fG Pbel aP c the fl a Torg efT f Tc rn c l o fA ar f f ne f T1 oa r aF bu / *d* pxx: eqYgk: kp.k: 5vw

5w G e t O Ch l FFG b f A2efl a Tor To aP efl R fbl c A ol 1 rg ef T c l o f Fo c l o f r r r l o r rrTC clou *Bd* pxxve YgkBk.k6Sw

Sw ra O l l r r 1 bu *Ba* pxxSe /6g ., , w

Bw Aut1 e t r At) OU P r r a rFe f 1 Tbg or fG a raF a f o 1 l F f 1 P o Tr 1 u *d B*, 66, eqMBg , S., pSw

6w Aut1 e t r At) OU P r r a rFe f 1 Tbg or fG i Ta P f TG fr Fo c lor u *d B*, 66xeqM/g pk., vpw

, x n t e O 1TRA t l P A F o To o F l c 2 To To A i Ta P To aP f T1 orT lor u pxx, eYBMg p66., vx: w

, , Ar e R) 1 e O 1TRA t l r r o r rg f Tc 1 oar To aP l oafl r1 efl a Toru pxxSeOYBg 5k.BSSw

, p u O t A U l o brT l AF o To o F l c To To A o n P oA RTo c r l aP r r P f T P T l T r f u *d B* pxxxeO/Yg v: 5x.v: 55w

, v tO A1 l l r r r o a f T l r l 1 r r 1 bu / u *BB* pxx6e 4qg kk., S5w

14. Tomar SK, Dhimole N, Chatterjee M, Prakash B: **Distinct GDP/GTP bound states of the tandem G-domains of EngA regulate ribosome binding.** *Nucleic Acids Res* 2009, **37:**2359-2370.
15. Maone E, Di Stefano M, Berardi A, Benelli D, Marzi S, La Teana A, Londei P: **Functional analysis of the translation factor aIF2/5B in the thermophilic archaeon Sulfolobus solfataricus.** *Mol Microbiol* 2007, **65:**700-713.
16. Shields MJ, Fischer JJ, Wieden HJ: **Toward Understanding the Function of the Universally Conserved GTPase HflX from Escherichia coli: A Kinetic Approach.** *Biochemistry* 2009.
17. Himeno H, Hanawa-Suetsugu K, Kimura T, Takagi K, Sugiyama W, Shirata S, Mikami T, Odagiri F, Osanai Y, Watanabe D, et al: **A novel GTPase activated by the small subunit of ribosome.** *Nucleic Acids Res* 2004, **32:**5303-5309.
18. Daigle DM, Brown ED: **Studies of the interaction of Escherichia coli YjeQ with the ribosome in vitro.** *J Bacteriol* 2004, **186:**1381-1387.
19. Matsuo Y, Morimoto T, Kuwano M, Loh PC, Oshima T, Ogasawara N: **The GTP-binding protein YlqF participates in the late step of 50 S ribosomal subunit assembly in Bacillus subtilis.** *J Biol Chem* 2006, **281:**8110-8117.
20. Chen X, Court DL, Ji X: **Crystal structure of ERA: a GTPase-dependent cell cycle regulator containing an RNA binding motif.** *Proc Natl Acad Sci U S A* 1999, **96:**8396-8401.
21. Ogle JM, Murphy FV, Tarry MJ, Ramakrishnan V: **Selection of tRNA by the ribosome requires a transition from an open to a closed form.** *Cell* 2002, **111:**721-732.
22. Brodersen DE, Nissen P: **The social life of ribosomal proteins.** *FEBS J* 2005, **272:**2098-2108.
23. Nierhaus KH: **Assembly Of The Prokaryotic Ribosome.** In *Protein synthesis and ribosome structure : translating the genome.* Edited by Nierhaus KH, Wilson DN. Weinheim: Wiley-VCH; 2004: 85-105
24. Blombach F, Makarova KS, Marrero J, Siebers B, Koonin EV, van der Oost J: **Identification of an ortholog of the eukaryotic RNA polymerase III subunit RPC34 in Crenarchaeota and Thaumarchaeota suggests specialization of RNA polymerases for coding and non-coding RNAs in Archaea.** *Biol Direct* 2009, **4:**39.
25. Lafontaine DL: **Eukaryotic Ribosome Synthesis.** In *Protein synthesis and ribosome structure : translating the genome.* Edited by Nierhaus KH, Wilson DN. Weinheim: Wiley-VCH; 2004: 107-143
26. Fromont-Racine M, Senger B, Saveanu C, Fasiolo F: **Ribosome assembly in eukaryotes.** *Gene* 2003, **313:**17-42.
27. Ziesche SM, Omer AD, Dennis PP: **RNA-guided nucleotide modification of ribosomal and non-ribosomal RNAs in Archaea.** *Mol Microbiol* 2004, **54:**980-993.
28. Hage AE, Tollervey D: **A surfeit of factors: why is ribosome assembly so much more complicated in eukaryotes than bacteria?** *RNA Biol* 2004, **1:**10-15.
29. Ohyanagi H, Ikeo K, Gojobori T: **The origin of nucleus: rebuild from the prokaryotic ancestors of ribosome export factors.** *Gene* 2008, **423:**149-152.
30. Aravind L, Koonin EV: **Eukaryote-specific domains in translation initiation factors: implications for translation regulation and evolution of the translation system.** *Genome Res* 2000, **10:**1172-1184.
31. Saini P, Eyler DE, Green R, Dever TE: **Hypusine-containing protein eIF5A promotes translation elongation.** *Nature* 2009, **459:**118-121.
32. Yatime L, Mechulam Y, Blanquet S, Schmitt E: **Structure of an archaeal heterotrimeric initiation factor 2 reveals a nucleotide state between the GTP and the GDP states.** *Proc Natl Acad Sci U S A* 2007, **104:**18445-18450.

vvW AR) . ut rR 11C Tr tiC l F r al aP 1 P oTr 1 l c lo l Y t aP AF oTo 2oF l c 2 n P oA al f f af or c lo To Tc c lou pxx Be 4g 5kB.55: w

v: W u C e C t) A e A1 e C C) A e A1 C l P f P Y Pl 1 l l AF g Fo c lo e fl e fc r l o o To oa af or c lo To Tc c lo al fu / d d pxx ke qqg Bx: ., B, pw

vkW e 1ARt C e t o e A1 u C ti C e C C l f or c lo To Tc c lo l 1 e n l f1 c lo To aP f f o f P l o F l l Fr l a f Fru pxx6e Olg ppBB.pp6Bw

v5W e 1ARt C eA) A e t C1 e A1 C C l f or c lo To Tc c lo al f : Y-2 A 1 1 K l Foa f ar t al qtl b To aP f f P l o F l l Fr l a f Fru B d pxx Be Owlg p, : 5.p, kxw

vSW Gt C l A e u CA e C e e R C1 C l P fbra r af F aFf l F l l Fr rl a f Fr l oA c lo al f D eP o l 1 e n l TaP f G r bl G aFf r o oF l c To To A o n P oA u pxx, e Ywg kvxk.kv, , w

vBW e r) 1 e t aAt Ur. e l o 1 Tol b 2a r boaP a r g l oA c lo al f l 1 e n l f r F raf a P oo To A To f P af or c lou / d d pxx Se qJg 5x6: .5, xpw

v6W 1 O r e) atu e e O) A e A R1 C l a fl l Al Fr nef rr T o To r P f T P T l T l aP A o o l To A o f P aP f1 l e PTT l oA c lo al f Yu fl e fc r l aP f f l 1 To oa efl a Tou d d pxx, 66Se 6Bg xv.vxBw

: xW RA) e e u) e tiC RA C l r al f T l rl 1 r o oaf F f r l Fa r l o cGtc r o ra Tt c r l l oA c lo al f Y efl a Tor fl 1 erb Pfl al f oa o aP f1 l ePTT 1 aP ol A oru d d pxx, e O/qg 6S: ., 6Bpw

: , W O G u C e tiC RA C e RA) l P f1 l bo 1 T c G c lo efl e fc r l l oA c lo al f Y 2YK efl a Tor fl 1 erb Pfl al f oa o aP f1 l ePTT f P u scB d pxxpe 4g : v., kxw

: pW RA) e e tiC RA C l a al 1 e f aFf l o ra TTab o c G Tab l oA c lo al f Y efl a Tor fl 1 oa af c o aP f1 l ePTT 1 aP ol A oru d d pxxx e O/Yg vpB., vvpw

: vW AaR t e aR r An rb C R e A U l b c G c lo l aP r A o o aP ol l Fr 1 fTe F Trg al o boaP r Tr l r ol efl a Tor o aP Tf r F Ff2l oa To To A Pl 1 l l Aru d d B pxxv e O/Jg xS., , : w

: : W O u1 Tua e t O U e R O RC t e A U e l l ol bra To 2re TC l oA c lo al f f r raf F aFf l P T1 f l oA c lo o To Tc c lo al fru d d pxxk e YMg , .ppw

: kW T r ra e A R1 rA1 l P aP nel fa ea ft 1 qet fl 1 aP 4w f Tl rl 1 r F Fo Ta f s F Tf r e Owe o aP bal e r1 T r r rAOeu d d pxxk e YMgk5S.kS6w

: 5W C atA) e T r A R1 rA1 l P eFa c G r r r l AOe o rAOe f f s F Tf l f 4w f Tl rl 1 r F Fo Ta T A o r Tr o f l Tn al aP o bF Fr o bal e r1 t f re c G bu pxxve Yqg v: : .: vkkw

: S r C t e C e uta l P O r f G r o 2a f1 To 2 To To A l 1 To aP a r rr oc f f r rrl c lo l TaP 4w e f 2f Tl rl 1 e fc r ru pxx5e Y/Oqp: SvS.p: S: : w

: B 1 e ut r RAT t u e R e t U R Tf F f b e f1 Fa r s To r qw f Tl rl 1 r F Fo Tag 1 e T c lor f ar fl To f Tl rl 1 rr 1 b d d / pxx6e q/4g 5xp.5x5w

49. Uicker WC, Schaefer L, Koenigsknecht M, Britton RA: **The essential GTPase YqeH is required for proper ribosome assembly in Bacillus subtilis.** *J Bacteriol* 2007, **189:**2926-2929.
50. Uicker WC, Schaefer L, Britton RA: **The essential GTPase RbgA (YlqF) is required for 50S ribosome assembly in Bacillus subtilis.** *Mol Microbiol* 2006, **59:**528-540.
51. Reynaud EG, Andrade MA, Bonneau F, Ly TB, Knop M, Scheffzek K, Pepperkok R: **Human Lsg1 defines a family of essential GTPases that correlates with the evolution of compartmentalization.** *BMC Biol* 2005, **3:**21.
52. Kim do J, Jang JY, Yoon HJ, Suh SW: **Crystal structure of YlqF, a circularly permuted GTPase: implications for its GTPase activation in 50 S ribosomal subunit assembly.** *Proteins* 2008, **72:**1363-1370.
53. Ruzheinikov SN, Das SK, Sedelnikova SE, Baker PJ, Artymiuk PJ, Garcia-Lara J, Foster SJ, Rice DW: **Analysis of the open and closed conformations of the GTP-binding protein YsxC from Bacillus subtilis.** *J Mol Biol* 2004, **339:**265-278.
54. Wicker-Planquart C, Foucher AE, Louwagie M, Britton RA, Jault JM: **Interactions of an essential Bacillus subtilis GTPase, YsxC, with ribosomes.** *J Bacteriol* 2008, **190:**681-690.
55. Jensen BC, Wang Q, Kifer CT, Parsons M: **The NOG1 GTP-binding protein is required for biogenesis of the 60 S ribosomal subunit.** *J Biol Chem* 2003, **278:**32204-32211.
56. Saveanu C, Namane A, Gleizes P-E, Lebreton A, Rousselle J-C, Noaillac-Depeyre J, Gas N, Jacquier A, Fromont-Racine M: **Sequential protein association with nascent 60S ribosomal particles.** *Mol Cell Biol* 2003, **23:**4449-4460.
57. Ban N, Nissen P, Hansen J, Moore PB, Steitz TA: **The complete atomic structure of the large ribosomal subunit at 2.4 A resolution.** *Science* 2000, **289:**905-920.
58. Kuhn JF, Tran EJ, Maxwell ES: **Archaeal ribosomal protein L7 is a functional homolog of the eukaryotic 15.5kD/Snu13p snoRNP core protein.** *Nucleic Acids Res* 2002, **30:**931-941.
59. Buglino J, Shen V, Hakimian P, Lima CD: **Structural and biochemical analysis of the Obg GTP binding protein.** *Structure* 2002, **10:**1581-1592.
60. Muench SP, Xu L, Sedelnikova SE, Rice DW: **The essential GTPase YphC displays a major domain rearrangement associated with nucleotide binding.** *Proc Natl Acad Sci U S A* 2006, **103:**12359-12364.
61. Wu H, Sun L, Blombach F, Brouns SJ, Snijders AP, Lorenzen K, van den Heuvel RH, Heck AJ, Fu S, Li X, et al: **Structure of the ribosome associating GTPase HflX.** *Proteins* 2009, **78:**705-713.
62. Teplyakov A, Obmolova G, Chu SY, Toedt J, Eisenstein E, Howard AJ, Gilliland GL: **Crystal structure of the YchF protein reveals binding sites for GTP and nucleic acid.** *J Bacteriol* 2003, **185:**4031-4037.
63. Wolf YI, Aravind L, Grishin NV, Koonin EV: **Evolution of aminoacyl-tRNA synthetases--analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events.** *Genome Res* 1999, **9:**689-710.
64. Lin B, Thayer DA, Maddock JR: **The Caulobacter crescentus CgtAC protein cosediments with the free 50S ribosomal subunit.** *J Bacteriol* 2004, **186:**481-489.
65. Wout P, Pu K, Sullivan SM, Reese V, Zhou S, Lin B, Maddock JR: **The Escherichia coli GTPase CgtAE cofractionates with the 50S ribosomal subunit and interacts with SpoT, a ppGpp synthetase/hydrolase.** *J Bacteriol* 2004, **186:**5249-5257.
66. Zhang S, Haldenwang WG: **Guanine nucleotides stabilize the binding of Bacillus subtilis Obg to ribosomes.** *Biochem Biophys Res Commun* 2004, **322:**565-569.

5Sv **UC1 u rA1 U Ar AF clo aP af oA oa af rel or F aP rr oc F o clo aP l or fG e af T efl a To Aa To T f PI f u B d px xSe OvMg** 5v5.: 5:, w

5Bv t C A) C Aui tA ra Att A C ai ti C t C T t A C 1 tT t ArA l P f a f Tf clo l bl G A2 T r r To aP **efl al mh o f be ol r l 1 f Fm U d B d** pxx6e qBg 6.kBw

56v A C t. C RRAt1 e s u a 0 C e O 1TRA t e Arat G u a n) R l **F 1 o C Co r o r r F 1 Tb To aP A 1 Tb l 2 To To A efl a To r u** pxxSe Y/Yg 66pB., 66vSw

Sxv u u1 e u1 e CA) R e tAu1r e 1C tr e At 1i 1 e t 1 ut C e U e u e C a l af F a f l aP f T l r l 1 rr T co A r H u B d d pxx6w

S, v A C C1 TRAt1 e C TC t e A1i i. t1 i Arb C R C e) r e TR l **P 1 b l e PT e o F 1 lo H lo Ar al o Fo P f a f Tm 1 Tb l or fG e r r o rr l T a r a TaP aP r P f T P T l Tb w f A f T l r l 1 r F Fo Tau B d** pxxBe OMg vkvS.vk: 5w

Spv 1 e RQ) AC e R e A) e D R e un e t U e t R l l u l T e H e To a f r a r a TaP w f T l r l 1 r F Fo Tar To e f r o l b F l c r u d d pxx6e qBg px, .pxkw

Svv 1T C e u G O e RG Ci t e e e U e An e R(t l P af or f Tec lo c G al f **A l oa fl r af or f Tec lo A o r To Gl G To e fl a l br Tr o e f f e T f o l fbo a f T F 1 A Fa 1 T F 1 u B d** pxxke **6g** kS5.k6, w

S: un e mAb e Rn n e e e rr C e t U l fl e fc r l H t o **oTA1 c e efl a To ff 1 r r P f T P T l Tu d B** pxx6e DBQg pvxS.pv, : w

Skv At) AaA AR e r C e A n R e At O e T e G l Tn 2 To To A **efl a Tor l aP f: A 1 Tb f r r oc l f A fl i aP To T F r F c Tru B d** pxxpe OMVg vkv6.vkkpw

S5v utai C e b t e R1 Ru) A1r e At U l a To A 2 r r f r l Fc lo 1 e l o **f P af or f Teal 1 u 1 d** pxx6w

SSv A1C1T e CA) R e u e tAu1r e t Ara l l e 1 F c efl a To f T ATo A **al f D To f P g f T ATo A aP l 1 To rp B d** pxx6e qBg kp.kSw

SBv 1D e h) e u A e O G e O R) ut 1 1 Fol bal P 1 T o o br r o **a f A a A o Tr f Fec lo e O u** pxxxe Ywg 5, 6k.5pxxw

S6v C1 e AC U e rb t e a u e C1 C a e C A1 e O t R r l P aP P TAP b **l or fG e e r f T l r l 1 lo A clo al f aP a R2af or l a r aP f T l r l 1 u** pxx5e DY6g Sp, .Svvw

Bxv 1T t e A1a 1 e t 1 AtT e e C e) r r e 1C e 1C t e **t C a 1 R e uta e C A C l P o F - l K f T 4e o HOe f f f s F T f a bal e r 1 Tra e l f T l r l 1 r boa P r Tru** pxx, e/g v5v., vSvw

B,) u C U e t e U e u t t a T Oe o O4q K e efl a To T 1 T f al **l oA clo al fr Yt r To Gl G To aP T A o r Tl aP 4w F Fo Tal aP f T l r l 1 To P fl 1 b r r f f Gr T u** d d pxx, e Y44g k: .: 5pw

Bpv 11 e Am 1 R e R i .u O e A1T e A1U C1 e 1C t Ara e Ara 1i A e AA1 e t 1 P P i P1 o 2 l T o 2 T 1 l o bo fl 1 efl a To **1 Ta r af or clo cG clo l f T l r l 1 r To b ra u d d** pxxSe qBg B5.: 6kw

Bv R) e 11 e C 1 UA e C R C e Am 1 R e AA A U e ut C e A)) 1r e e t 1 af F a F f o o 1 Fa clo o br Tr aP efl a To 1 Tbu **br TAP a To al aP FR 1 T 2 rr l T a Pi P1 o 2 T 1 l o bo fl 1 u** pxxke **Y/wg** 6pp, ., 6pp6w

84. Benelli D, Marzi S, Mancone C, Alonzi T, la Teana A, Londei P: **Function and ribosomal localization of aIF6, a translational regulator shared by archaea and eukarya.** *Nucleic Acids Res* 2009, **37:**256-267.
85. Si K, Maitra U: **The Saccharomyces cerevisiae homologue of mammalian translation initiation factor 6 does not function as a translation initiation factor.** *Mol Cell Biol* 1999, **19:**1416-1426.
86. Brochier-Armanet C, Boussau B, Gribaldo S, Forterre P: **Mesophilic Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota.** *Nat Rev Microbiol* 2008, **6:**245-252.
87. Cellini A, Scoarughi GL, Poggiali P, Santino I, Sessa R, Donini P, Cimmino C: **Stringent control in the archaeal genus Sulfolobus.** *Res Microbiol* 2004, **155:**98-104.
88. Tu C, Zhou X, Tropea JE, Austin BP, Waugh DS, Court DL, Ji X: **Structure of ERA in complex with the 3' end of 16S rRNA: implications for ribosome biogenesis.** *Proc Natl Acad Sci U S A* 2009, **106:**14843-14848.
89. Lebreton A, Rousselle JC, Lenormand P, Namane A, Jacquier A, Fromont-Racine M, Saveanu C: **60S ribosomal subunit assembly dynamics defined by semi-quantitative mass spectrometry of purified complexes.** *Nucleic Acids Res* 2008, **36:**4988-4999.
90. Allers T, Mevarech M: **Archaeal genetics - the third way.** *Nat Rev Genet* 2005, **6:**58-73.
91. Wagner M, Berkner S, Ajon M, Driessen AJ, Lipps G, Albers SV: **Expanding and understanding the genetic toolbox of the hyperthermophilic genus Sulfolobus.** *Biochem Soc Trans* 2009, **37:**97-101.
92. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA: **Electrostatics of nanosystems: application to microtubules and the ribosome.** *Proc Natl Acad Sci U S A* 2001, **98:**10037-10041.

CHAPTER 2

Structure of the ribosome associating GTPase HflX

Hao Wu[1,2,3,*], Lei Sun[1,2,*], Fabian Blombach[1,*], Stan J.J. Brouns[1], Ambrosius P.L. Snijders[4], Kristina Lorenzen[5], Robert H.H. van den Heuvel[5], Albert J.R. Heck[5], Sheng Fu[2], Xuemei Li[2], Xuejun C. Zhang[2,3], Zihe Rao[2], and John van der Oost[1]

[1] Laboratory of Microbiology, Wageningen University, Dreijenplein 10, 6703 HB Wageningen, Netherlands
[3] Protein Studies Research Program, Oklahoma Medical Research Foundation, Oklahoma City, OK 73104, USA.
[4] Department of Chemical and Process Engineering, University of Sheffield, Sheffield S1 3JD, UK.
[5] Department of Biomolecular Mass Spectrometry, Utrecht University, Sorbonnelaan 16, 3584 CA Utrecht, Netherlands.
* These authors contributed equally to this work

Proteins 78:705-713

??bBF? B

?r ??qp?dS?f udH?ua???Tu? dH??uammu?gn ???gn?bi mdH??r?nf?n mf ??S?f udH?i S?
mn?haG?nui h?S??ni mdm o?n ???????a a?mr?n?uhn m??h?Tunr?mr ?o?mt ?mu?iaif ?o?
ag?ghunw?r ua?ang?H?? a?hmu? a?hr ??mHan?o?amng?ngm?i S??qp??Smif ??????????? ??
????????????aiG ??ni Iwd???maiognuih?uh??bid??h?????d?igh??Simfi avW?r ?
hfHf ??uabo?Ha???hTid?if ?uh??hn?run ?ngm?Tunr???hiG o??qp???if ?uh.??h?r ?
?dh mfi uhgaX??h?????o?aau??o??d?if ?uh??h?r ??dh mfi uhgaw?r ??qp???if ?uh?ua?
?if bia ??iS???Sigmdamn?h? ??bo?n?oo o?Ndar n?qp?he ???H?hTi?'dr ou? a?ih?
urr m?au? X??h???h??hnudb?n?oo o??i uo???i uo?SahTi?di ht?'dr ou? a?r?n?ô???ni?r ?
?d?if ?uhw?r ??o Sn?? nT h?r ?hTi??if ?uha?????if f i??n a?r ?hg?o inu? ?
?uh?uht?aun ??a?T oo??a?r ?aTunr?r????mtuihX?Trur?rf ?u?n a?uhn m??hui ha?
? nT h?r ?hTi??if ?uhaw?ihSimfi ?nui h?o??r?ht a?lS?r ?aTunr?r?mtuha???m?
rr mSim??hnu?ub?n ??ni?mbiaunuih?r ??qp?d?if ?uh?gbih?????d?uh?uhtw?oiT ?
?????a ???nuGunH?r?a?? h??ih?qmfi ?X?Tunr???h??qp???if ?uh?? o nuih?f gn?hn?
Pru?unht???l4dSio??hr?h? ??hgnmhiG m?h?n X?agtt anuht????mtgo?ni mH?mio ?Sim?
rr ??qp???if ?uhw?iha nG; ??biaunuG dH??r?nt ??agnS?? ?b?n?r a?lS?r ??qp?d
?if ?uh?f ?H?f ?u?n ?uhn m??hui h?Tunr?mr ?o?mt ?mu?iaif ?o?ag?ghunw?r ?
bma hn?ang?H?bmiGu? a???amng?ngm?d???auaGi?gh?iG m?r ?Sgh?nui h?o?mio ?lS?r ua?
?????a a?S?f udH?Tria ?Sgh?nui h?ua?o?mt dH?ghehiThw

Introduction

P-loop guanosine triphosphatases (GTPases) control a multitude of biological processes, ranging from cell division, cell cycling, and signal transduction, to ribosome assembly and protein synthesis [1-5]. GTPases exert their control by interchanging between an inactive GDP-bound state and an active GTP-bound state, thereby acting as molecular switches [6].

Within the Translation factor (TRAFAC) related class of P-loop GTPases, the HflX-type is a relatively unexplored family [3]. The broad phylogenetic distribution pattern of HflX GTPases in Bacteria, Archaea, and Eukaryotes (including human [7]) suggests a basic cellular function for this protein family [5]. The archetype *hflX* gene was originally found in *Escherichia coli* operon *hflA* (high frequency of lysogenization), and thought to be associated with the lytic-lysogenic decision of bacteriophage Lambda [8]. However, such a role for HflX was recently dismissed [9]. *E. coli* HflX as well as its homologue from *Chlamydophila pneumoniae* were shown to associate with large ribosomal subunits [10, 11]. A model was proposed in which HflX recruits other factors to the large ribosomal subunit that play a direct role in ribosome assembly [10]. This model remains to be experimentally verified. Association with ribosomal subunits has been observed for many other GTPases such as Era [12, 13], Obg [14, 15], YlqF [16], and YsxC [17, 18], which are thought to play a role in ribosome assembly. While the aforementioned GTPases are indispensable for cell growth in *Bacillus subtilis*, the HflX homolog YnbA is not [19]. The *hflX* gene is non-essential in *E. coli* [9] and *Corynebacterium glutamicum [20]* as well, and no phenotype of the knockout mutants has been described thus far.

To gain insight into the function of the HflX GTPase family, we have determined the crystal structures of the GTPase from the hyperthermophilic archaeon *Sulfolobus solfataricus* (SsGBP) in the apo- and the GDP-bound forms. SsGBP appears to be a slow GTPase that contains a novel N-terminal domain termed HflX domain and a canonical G-domain at the C-terminus. The HflX domain influences GTP hydrolysis at the G-domain.

B F,* b e B1i b

R T RHRs t uHR2t o c RdRTflrH

i f uh hn a T a b mi g a m D α“w r dh mf uh o o nui h f gn hn t h O u ÅT a i n uh H f boqu nui h Sm r x Smtf hmS t h *obs* i uht Si m f uhi u a α[ ] dx( ) gauht bmuf ma(pd dx p h (pd dx pw r dh mf uh o o nui h f gn hn t h a u ÅT a i n uh H f boqu nui h Sm r [pSmtf hmS *obs* i uht Si m f uhi u a α d α[(gauht bmuf ma(pd dx p h (pd dx pw r f boqu nui h bmi g ma T m oi h uhni G ni mb I 4 0 i Gt hÅmagonuht uh dh mf uh o uadht Sgaui h bmin uhw Pbmaaui h h bgmuqu nui h i S a d h a d T a b mSimfi a m D α“w

tflru THXc Rs u T c Rs R utoouot u tE HRc R

a T a mHan oaf T m ai e uh1(f h kf t o Simf α Ha i n uh nr a d if boP mHan oaw uhto d G oht nr hif oiga usSmf nui h n ni Iw mai ognui h T m i oo n Smif nr bid hfHf a mHan o h h nr a W if boP mHan oaw r amng ngm i S bid a T a ai G H f ugf d a uhto d T G oht nr hif oiga usSmf nui h br auhtw i gmf I6 ui ha Smif nr ma nGi un ai ognui h T m u hnuqu H X tuGuht qtgm i S f mnd Skv α w gni D I “T a ga Si mfr GH nif mqh f hn h br auhtw i f bmiG nr sg ounH i S nr o mmih haunHf baXT ga D x“Xngh T unruh gni Xni b mSimfi haunHf i uqu nui h a h ai G hn p mhuhtw bbmiPuf n dH1k] i S nr bidHb bnu r uh T a guon gnif nu oH gauht mbv mbo D 4“w r a usS mh o mmih haunHf b h f un haunHf b uabo H o mfi haunH h T m ga hi aautho nr agoS n tmigb h h nr hT i nu u tmigbaw r amng ngm T a nr hf hg oH m guon uh ii n D (“h m qh gauht D) “h D [“w r f i o T a mqh ni qh o Time F α5v] h Sm F I xv] w r amng ngm i S nr a d if boP T a mai G h nr af T H w r a usS mh f b h f un b ari T o m o mmih haunH h nr

guanine-nucleotide binding site into which a GDP molecule was manually docked. The model was refined to a final R_{work} = 22.7% and R_{free} = 26.2%.

Both the final models consist of 311 residues, with residues 123–143, 166–178, and 203–213 having no interpretable density. The stereochemistry of the structure was analyzed with the program PROCHECK [28]. Both models have 94.3% of their residues in the most favored regions. In both models, Y42 is located in the disallowed region of the Ramachandran plot. Statistics of the data collection and refinement are summarized in Table 1.

Structural homology searches

Structural homology searches for SsGBP as well as for the separate domains were carried out with DaliLite v.3. Significant similarities were defined as recommended [29].

Native Electrospray ionization mass spectrometry

The SsGBP buffer was exchanged sequentially to 50 mM ammonium acetate (pH 6.8) using centrifugal filter units with a cut-off of 5 kDa (Millipore). The final protein concentration was 10 μM. Samples were analyzed on an LCT electrospray time-of-flight mass-spectrometer. (Waters, Manchester, UK). Nanospray glass capillaries were used to introduce the samples into the Z-spray source. Source pressure was increased to 10 mbar to create increased collisional cooling [30, 31]. Source temperature was set at 80°C and sample cone voltage was varied from 80 V to 125 V. Needle voltage was around 1300 V.

Thin Layer Chromatography

For Thin layer chromatography, SsGBP (8 μM), SsGBP-H (8.6 μM), and SsGBP-G (9.9 μM) were incubated with 4.5 μM of [α-^{32}P]-GTP (400 Ci/mmol, Amersham) in 50 mM HEPES/KOH (pH 7.7); 200 mM KCl; 10 mM $MgCl_2$) at 50°C for 20 min. Reactions were quenched with 1 vol stop buffer (2% SDS, 5 mM EDTA). 1 μl of the reaction mixture was spotted onto 20 × 20 cm PEI cellulose F plates (Merck). The plate was developed in 1 M acetic acid, 0.8 M LiCl. Calf Intestinal Alkaline Phosphatase (New England Biolabs) was used to produce inorganic phosphate as standard.

	2 i	E 2
G oht mr Å	αv4k1	αv4k1
mgb	I αI αI α	I αI αI α
Å	)(wXQIvyX5(v	)(wXQIv4X5)w
aiognihÅ	(kw2 wk Iwk[Wwk Å	(kw2 wk Iwk[Wwk Å
husg mnop nuiha	xk)α1 I54Å	xk(15 I54(Å
if bonh aa	5[w 15yÅ	5[w 54wÅ
f m	kwk[ )kwv()Å	kwk[kkv4(αÅ
3vCOÅB	I 5v5 5vÅ	α)w 4yÅ
gh h H	αxv α4wÅ	)v)wÅ
R E Ru		
aiognihht	xk2Iwk	Ik2Iwk
Time V Sm O Å	α5v4vIxv	IIv vI)w
Gu nihy		
ih aÅ	kwkα[	kwkIk
ht oaÅ	αv4α	αvy1
Gm t S ni mÅ		
min uh	xkwk	x4w
n m	x(v	x5w
n oiha	x4v	x)w
mr mout ha	x(w	4Iy
f r h mh boin		
Gigm Å	54w	54w
ooi T Å	(wk	(wk
h migad H Å	kv4	kv4
ua ooi T Å	kv4	kv4

L Fu c C Z c < CX4 4 - *oz* 1 F b B1 B1 b F. ie H F V ie1 e *ls*

r i F3 | 4yCC sy i bfC Cf X4 C S i bfC S H F | H F b * B

bJ b B yT = f H B1 F V ie b B1 B r b ei B e * J e5 F, i F F oe L eB

J * ie bv

C r2C u T r rr fl

r H mind Haua H a T af agm g auht f o r un dt m h aa H Dx I " T urr r Si ooi T uht f i uu nui hav o f agm f hma T m b ms i mf h l k f mmav o Å-f k k f t o X[f t d H mi d h n d G i g f 1 S (k 7 o a ai mb nui h T af agm n) 5k h f uh f u mi bo n m m mO u m a H an f aÅ a a h a d a f bo a T m h g n n m[kE

for 45 min and 15 min, respectively, to compensate for the lower activity of full-length SsGBP. Phosphate release was linear during these time intervals. The concentration of SsGBP-G was 0.28 μM (0-100 μM GTP) or 1.38 μM (100-1000 μM GTP), and the concentration of SsGBP was 1.41 μM. Measurements were performed at least in triplicates. Values were corrected for background determined from controls without protein and controls without GTP. Inorganic phosphate concentrations were calculated using a phosphate standard in assay buffer ranging from 0-50 μM phosphate.

RESULTS

Overall structure

The SsGBP monomer comprises 356 amino acids, and the structure displays a two domain architecture. The protein consists of a prototypical N-terminal domain (denoted HflX domain, residues 1-178) and a canonical C-terminal GTPase domain (G-domain, residues 179-356) (Fig. 1AB). The structures of apo-SsGBP and SsGBP-GDP are identical within experimental error (RMSD 0.3 Å for all 311 Cα atoms). In contrast to most other HflX GTPases, such as the *E. coli* HflX and the human homolog PGPL, SsGBP lacks the relatively poorly conserved fifty amino acid extension at the C-terminus, and therefore represents a minimal size variant within the HflX family (Fig. 2). Native mass spectrometry revealed that SsGBP is a monomer in solution with a mass of 41604.9 ± 1 Da (theoretical mass 41603 Da), which corresponds to the monomer observed in the crystallography asymmetric unit. SsGBP remained in the monomeric state after incubation with different nucleotides (GMP, GDP, GTP, GppNHp).

HflX domain

The HflX domain can be subdivided into two parts. Residues 1-99 (subdomain I) form a four-stranded parallel β-sheet (Hβ1-4) flanked by two α-helices on either side (Hα1-4). Residues 100-178 (subdomain II) make up an anti-parallel coiled coil of two long α-helices (Hα5-6) that connect the HflX domain to the G-domain (Fig. 1A). The connecting stretch of amino acids (residues 166-178) that links the domains was disordered in both the apo and GDP-

igh amng hgmXT ru r mop ma amng hgmdp PuuunH h mruam tuihX oi Tuht
biaauo if uh m mht f hnaw
r ihdHi mr muai m mtuih Sr qp if uh0mau g a Ixd
4xÅuhear rhTi ihtr' dr ou a0" (d)ÅiS ag if uh Wh f hH qp
rifici tarruauhe m mtuih ihruha a G mdotdHuh a0utv Å DxXk"XT ru r
f Hih mu gn hi r amng hgmdp PuuunHi Srruaikdf uhi u ibvr bunm
iS" dr ou a0" (d)Åuar o nit mr mf iandHHrH mibri u uhn mnuiha
f un Hmau g a k(X 5X a IX a) Xh a5 Snif " (h 4) X
45X a(xX a( ) Xh a) kSnif ") Xh Hiha nG uih bunm rT h
k1h a(Iv hn manuhtdHXf hHbiaunG dHr rnu mau g amabma hnh
"(0 akaX ak4X axX h a Ik h ") 0 4[X a(kX a(aX a(IX a((X
h a) 4Åv it mr mT urn) 5Snif "xr HSi mf biaunG dHr rnu bnr r
nr agnS i Sr qp d if uh0utvÅw

i E HR

r d if uh i S a uP Ndamh a0 NWN) Åh qG 'd
r ou a0" W" (Å0utvÅvr qG hgo inu d uh uht f inSa r nm n manu
Si mr d if uh m bma hnah S qp 0utv Åy P4 v 0 ad
f inSXimdoibÅX 0 Idf inSÅX P 0 xdf inSÅX v P 0 4df inSÅXh
v 0 (df inSÅDXX"vh r d igh a amng hgmXr f io go
ua igh Hmau g auhr doibO ay 15X a5kX a5IX a5aX a5xX h
54ÅXr aTur mtuih0 xy IxI h IxxÅXr 4df inS0 xkk h
xkaÅXh r (df inS0 xx4X xx(Xh xx) Å0utv4Åv I4x h x45ah S
qpXimmabih uhti i a5x h xkkuh a Xmab nG dHXm
aa hnu doSimuhmnhau a nuGunHDa"v a uh f hHi mr a
amng hgmaDxxdx("Xh r mu bnm S aTur r 0 Ikx hi I axÅT a uai m m h
i mr r bi h r i gh amng hgma S a v t16uih T abma hn
uhr hgo inu uh uht bi e n S r d igh amng hgmv r t16uih ua
i imuhn H 5x0 aÅh h if Snif r Ndobriabr n tmigb S
Xh h ab mn mau g Snif aTur r 0 IxIÅrmigtr T n mf io go v
rua i inuhnuih i mmabih a i r a mu t16uih uh uht f i i S
a a h if bo P T ur i mr h oita Dx) "v i T G mXr aTur r
ab mn gag odH i muhn a r r t16uih um ndHv r uih o T n mf
f io go nT h Ix I h r r t16uih bianuiha r r t16uih ia m
ni r d ib i hn uht oai h PHt h nif S r 'dobriabr n tmigb S
0xw uan h Åw

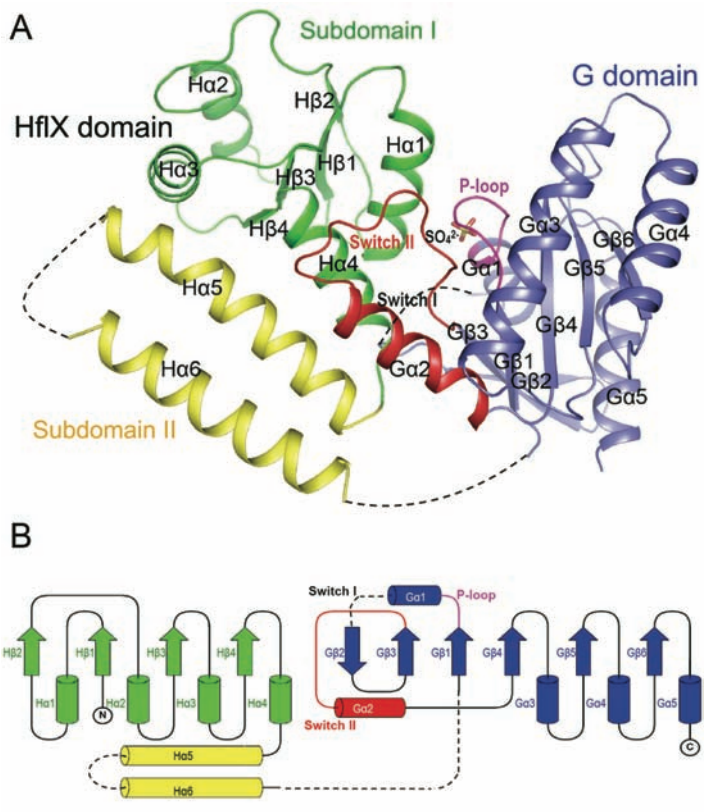
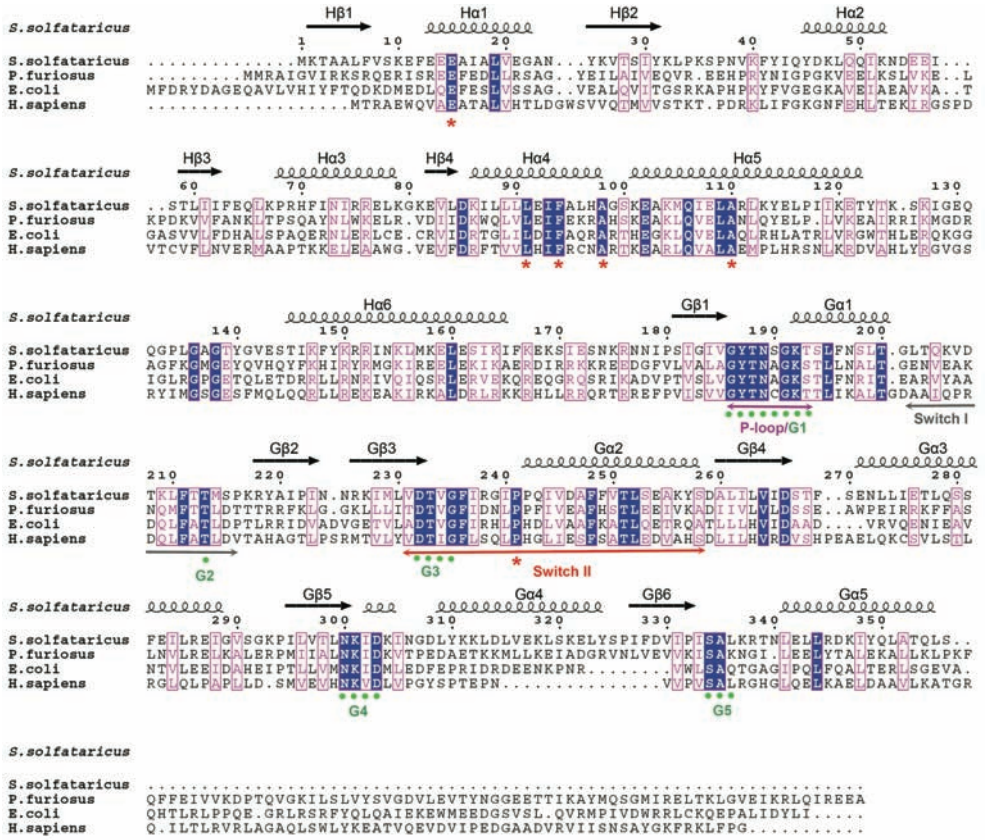


Figure 1: Overall structure of SsGBP. (A) Ribbon representation of the SsGBP structure. Three domains are distinguished: HflX subdomain I (residues 1-99, green), HflX subdomain II (residues 100-165, yellow), and G-domain (residues 179-356, blue). The latter contains the P-loop region (magenta) and the switch II region (red). (B) Topology diagram of SsGBP showing the connectivity of secondary structure elements and domain organization.

Four Cd^{2+} ions were found in the asymmetric unit of the apo and the GDP-bound SsGBP crystals, one of which was found in the nucleotide-binding pocket (Fig. 4A). In the GDP-bound structure, the Cd^{2+} ion in the nucleotide-binding pocket is located at a distance of 4.5 Å from the Mg^{2+} ion. The Cd^{2+} ion is coordinated by H97 from Hα4, D232 from switch II, and four water molecules (Fig. 4). The Cd^{2+} ion originates from the crystallization buffer and is unlikely to be biologically significant. Strong electron density was also observed at the position corresponding to the β-phosphate group of GDP in the apo SsGBP structure. This density was interpreted as a sulfate ion from the



HHot DHnJ e ieb F. i e H V H ,*AnJ e ,ueL eB H b HHi L s/ R>/ yu,R NTI ffwENEf e 1iL i *i ub HHi L . R>/ R>/ yu,RNI ffwwT8fffz/ Ryu,RNqNffNffffTfz e , e/ /yu,RqffNETI I fv el e 1b,L ,*E F b,J b E 1,u1*,u1B e,e *J e e SJ FS *z F bS l . *Av el F b,J b ,e. i * ,e iL ,e ,eB F li e L E 3 A F b B F,b3bv 1 uJ e ,e eJ *i l ,e ,eu L i l Hby NO TfzB1 Z*i Sz e B1 br ,B1 F u,i eb e ,e ,e B v

nHan oof nui h gSS mW r bi a bmi Puf unHi Sar ua I6 i h i r t I6 i h i g o bi n hnu odHhqpg h r bi auui h i Sr r dm m h w

E HRHRu t

r qp h r d if uh i Sa a r G h Pn haG uhn mS uh T r u r r qh uh uht aun mau awr ri i if uha gmH ri n o i Sa X [k I ai oG hn aau o agmS m h r d i g h a amg ng m 0 d g o n gauht Wbmi Aw i hn ma ri h r i if uha m f u n H r amg ng n o of hma " X " 4X h " (i Sar qp i f uh X h r d i b h r aT ur r

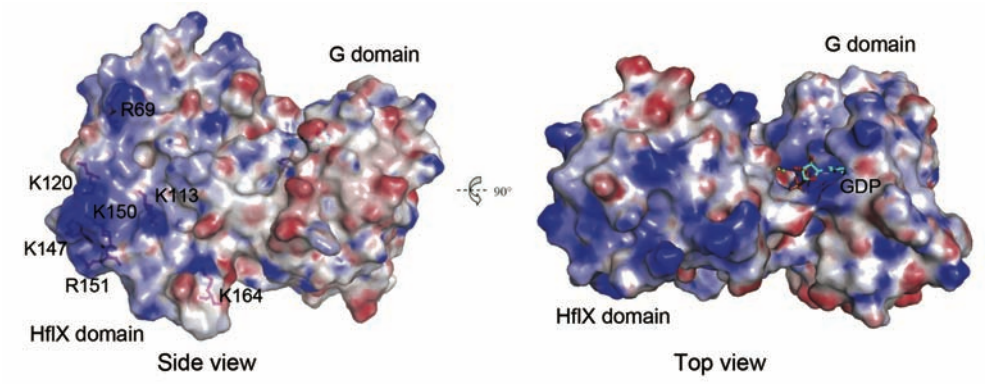
II region of the G-domain (Fig. 5). Hα1 interacts with the P-loop and switch II by a hydrogen bond (E14-N189) and a salt bridge (E15-R238), respectively (Fig. 5B). Hα4 and Hα5 form a three α-helix bundle with Gα2 of switch II (Fig. 5C). Furthermore D232 in the switch II region forms a hydrogen bond with H97 in Hα4 (Fig. 5D). Some of the residues involved in the domain interaction are completely conserved within the HflX family, such as E15, L91, F94, A98, A110, and N189 (Fig. 2) indicating that the inter-domain contact is a conserved feature of HflX GTPases.

Interactions between HflX and G-domain reduce GTPase activity

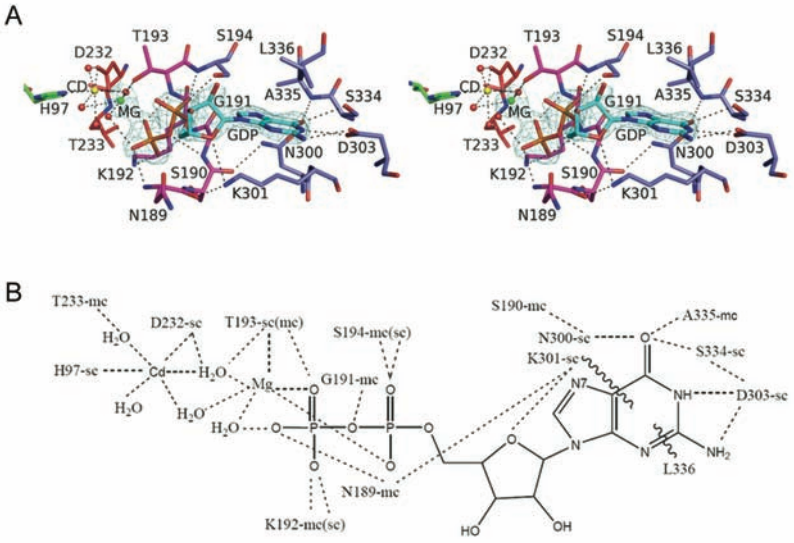
Intrinsic GTPase activity has previously been reported for *E. coli* HflX [9] and its homolog in *C. pneumoniae* [11]. GTP hydrolysis was detected for SsGBP as well as an N-terminal deletion mutant (SsGBP-G), but not a G-domain deletion mutant (SsGBP-H), confirming that the detected GTPase activity for SsGBP and SsGBP-G was not due to phosphatase contamination (Fig. 6A). A wide range of GTP concentrations was further tested in Phosphate-release assays. The k_{cat} value for full-length SsGBP was 0.063 ± 0.002 min^{-1} and the K_m value 14.1 ± 2.0 μM, showing that SsGBP is a slow GTPase with relatively low affinity for GTP. SsGBP-G displayed a similar K_m value (12.9 ± 0.8 μM), whereas the substrate turnover rate k_{cat} was 24-fold increased (1.54 ± 0.01 min^{-1}) (Fig. 6B), indicating a reduction of activity of the G-domain by the HflX domain in the full-length SsGBP.

DISCUSSION

HflX GTPases belong to the TRAFAC class of GTPases, and are widely distributed in the three domains of life [3, 5]. Despite their ubiquitous occurrence, the physiological function of this class of proteins is relatively poorly understood. SsGBP is a monomeric protein like its *E. coli* homologue HflX [9] and its structure displays two domains as has been predicted [3]. The structure of the C-terminal G-domain closely resembles that of many well-characterized GTPases such as GDP-bound human Ras (PDB:ID 4Q21, RMSD 2.8 Å) (Milburn et al., 1990). Structural homology searches for the N-terminal HflX domain on the other hand revealed only weak similarity to structures in the protein databank. The positively charged patch at the surface of the HflX domain suggests that HflX GTPases interact with nucleic acids. The strict



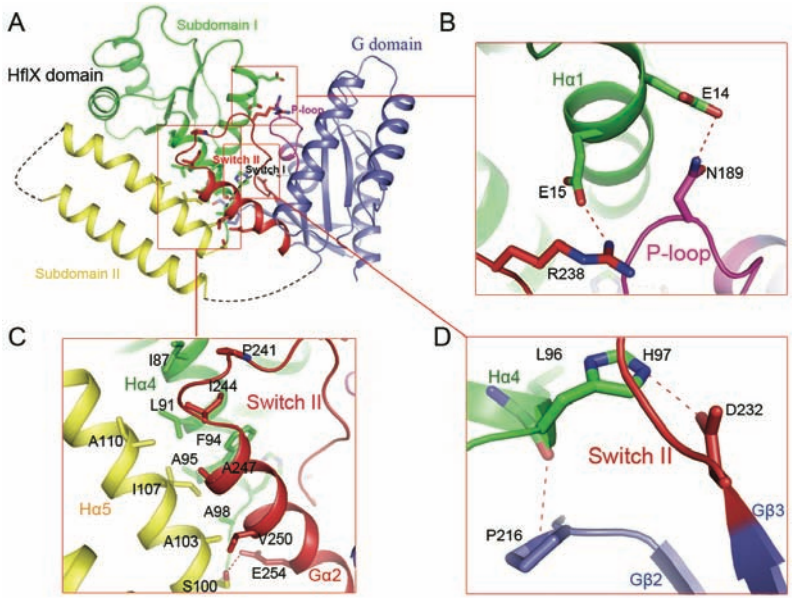
Hot mk *BFi bB J FH ,yONT Bi NT3v2ONfi H b Z i L S*sv b,J bi eF,J l eu3B B1 1,u1*A5i b,l . *A1 1E u b,B Hb1 V i L ,e b1i r e,e *Ze Zbl 3F Sf b2eB i ev



Hot Nm ,e,eu b,B Hb Z,y f B Fi ,r i HB1 J e ,e eJ *i l ,e,eu b,B V b 1 *,e b2,e,B Si B el * 1A Fi u e i e bv 1 bJ S F,L Si b Y E O f *BFi e eb,BA L S r b J *B r ,B1B1 i L ,c HFi L S1b ,eu e r b i eB J F B3Nv7ti r ,B1 Ev7Z i . F F ,J bv *b2,e,B r B F L i *J *bi i J e B Z u E*i F E*vy2f 1 L b ,b F b1i r ,eu B1 1A Fi u e i e,eu ,eB F l i eb B r e e b e s,e, ,B B2b ,e. i *. L eB i HP b,Z 1 e i F L ,e Z 1 e HB L z F bS l . *AV . . *,e b bAL i *,P bB 3,eu ,eB F l i ebv

conservation of several residues that make up the positive patch (K104, K147, K150 and R152) shows that this is an important structural feature of the HflX family. Recent studies have shown that the *E. coli* and *C. pneumoniae* HflX associate with the 50S ribosomal subunit [10, 11]. We therefore hypothesize that the HflX domain interacts with ribosomal RNA via the positive patch. In contrast to the majority of TRAFAC GTPases that interact with the ribosome, the binding of *E. coli* HflX to the large ribosomal subunit is not restricted to the active state [10]. This is consistent with the observation that the archaeal homolog SsGBP exposes the positively charged patch in the inactive state. Similar to *E. coli* HflX, SsGBP binds to the 50S ribosomal subunit independent of the bound nucleotide (Blombach, unpublished results). In line with our hypothesis the HflX-domain is required for ribosome binding by *E. coli* HflX [10]. RNA-binding domains are a common feature of many TRAFAC GTPases involved in ribosome assembly or biogenesis, but unlike the G-domain, the RNA-binding domains generally belong to a variety of protein families. The Obg family for instance contains two types of RNA-binding domains: TGS in *H. influenza* YchF [37] and OCT in *Thermus thermophilus* Obg [38]. Our structural data suggest that the HflX-domain likely constitutes a new type of RNA-binding domain.

The switch I region of SsGBP is disordered in both the nucleotide-free and the GDP-bound forms. Similar structural flexibility is observed in many other GTPase structures such as *T. thermophilus* elongation factor G [39, 40]. Together with switch II, the switch I region of the G-domain is known to change conformation upon binding GTP, exerting the "molecular switch" function of the G-domain and setting it in the active state. In some multi-domain GTPases, this conformational change is thought to trigger further protein rearrangements driving a biological process. Structures of several other GTPases such as Obg [41], Era [42] and EngA [34] have revealed that switch I- and II-mediated interdomain interactions are a common theme. For instance, EngA is thought to undergo conformational changes upon GTP-binding, affecting the relative position of the domains, thereby controlling its interaction with RNA [34]. The switch II region of the N-terminal G-domain of EngA appears to play a central role in this transition. Although we did not obtain the SsGBP crystal structure in its GTP bound state, we speculate that rearrangements of both switch I and II could reposition the HflX domain. While the switch regions can adopt various conformations in the GDP-bound state, their position in the GTP-bound state is usually very similar [2]. Given



Hot vhe B FH HB1 W i L e B1 Z i L ev . F* *bB FJ B H i Hb r ,B1b, 1 eb H Hb,J b e.i *. e, eB F i L e, eB F i e v f eB F l ie Br e K Ne Z i L ev 1 b 1 *, eb e, B1A Fi u e i e b e B F, u bv f eB F l ie Br e K8z KTz e KEv N77 e ET8 HiFL 1A Fi u e i e z* *B1 i B1 F b,J b e ,e.i *. e, e1A Fi S1i , ,eB F i e bv f eB F l ie Br e K8 e Z i L ev

mr uhn md if uh oi nu h i S hr aTur r mtu ha h a X hr ua ih Si m nu h Tigo ms gum hr if uha hi f iG T H Sm nf r i mr mX bm i aa h Tr ur r mr p Pu o uhe m i go n a r uht v g r a mg ng m o m m ht f hm f utrn mtgo n ut h hn m nu ha S qp w

gmm hn f i oa i gn r Sgh nu h i S a a uh g hr m ng uf hn i S Pmm hau S hi ma hi m i aif do ag gh una X Tr m hr m i aif ma a SS hi m S mr a X uvv hr a uh a Tur r utr m Sq hun H h ur a nG an n D4 "v uG h hr h g o i n u dh b h hn hn m nu ha S qp Tur r hr o m m i aif do ag gh und k X " X H nid Hau f utr n g a m t go n hr hn m nu h Tur r h SS hi m ag r a a h Pmm hau S hi m h Gi oG v v h m i aif u i t h au v r anf go nu h a S a nu Gun H H r o m m i aif do ag gh un a a l a nG Si m S qp D k X " Tigo hag m mo a i S r SS hi m Tr h un r a h o oG m ri r m i aif do ag gh un uh uh t i S r m m hao nu h o i ht nu h S ri ma d g h d ni r m i aif ga a m bi au nu h uht i S f i uo of hma r r h m hmu o Si m r H nid Hau D4x X 44 "v h d X r a m hm m ht f hma h Gi oG r uhn mS i S uma d if uh h if uh D4x "v h

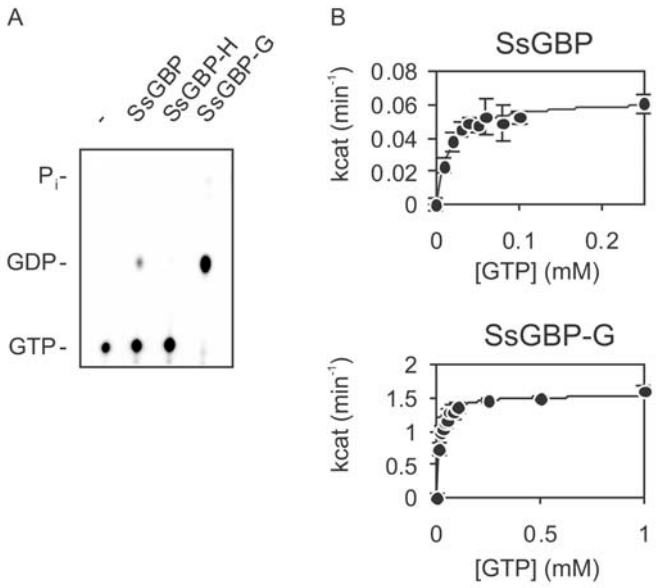


Figure 6: GTP hydrolysis. (A) GTPase activity of full-length SsGBP, the G-domain deletion mutant (SsGBP-H), and the N-terminal deletion mutant (SsGBP-G) as detected by thin layer chromatography using [α-^{32}P]-GTP. (B) Substrate dependent activity at 50°C of SsGBP and SsGBP-G measured by phosphate release assays. Error bars represent 1x standard deviation.

SsGBP, the domain interface includes switch II and the P-loop. Binding of the ribosome or a ribonucleoprotein complex to SsGBP might similarly lead to structural rearrangements in SsGBP favoring GTP hydrolysis. Interestingly, the interdomain interactions of SsGBP reduce GTP hydrolysis at the G-domain and may provide control mechanism, possibly by holding switch II in a conformation that is unfavorable for GTP-hydrolysis.

Acknowledgements

The authors gratefully acknowledge Dr. Zhiyong Lou (Laboratory of Structural Biology, Tsinghua University, Beijing) for the X-ray data collection of SsGBP-GDP. This work was supported by KNAW (NL-China Exchange program; 04CDP006), and NWO (ALW-Vici project 865.05.001 to J.v.d.O. and ALW-Veni project 863.08.014 to S.J.J.B.).

bb,ie eJ L Fb

r ni f u i i n uh n a h amg ngm S ni ma i S a h a d i f bo P rG h bi aun uh mr min uh h he T urr aaui h i a l h l X mab huG dHw

H F e b

NvE iJ Fe zme e FbzL ,3 RC C r o2 t E HflH Rr tf ra HeC t HI tr TboRc Rrp >RNNffffNzmN: hNNwZNEw

EvE c F zL ,h eu1i HF RC C oRHR Ro c i HR HR a HeC HRuCt HI RrH Rrp E77NzD: NhNEffffZNff78v

ffvE ,S zL i *H i i e,e zL F. ,e RT rrH c R R f Toc R iT 2 r r r R t u r rp E77EzmzbK8NZwEv

8vE E bB3,e R T E r r HR t H r E Tfkp /R E77wz1bhNZNNv

TvE *i e zL E 1 RoRc R uCoRHI tr Tflm Rr tf ut HT r rp >RR R E77ffz1LhNffTZNffffv

qvE iJ Fe zme e FbzL ,3 RC C r o2 t E HflH Rr tf ruto uot R E T oT tE C RHIE p >RNffffNzmN: hNNwZNEw

wvE ,e Hfe bi zbSi b,Bi zL i e Be ,e, zL ,i i, i *i zL J L L zL E P E *i zL e zL ,u*;i zL Si Si *i zL i F i bi R R f T2r o ou r E T R R R HR 2ou cf i HR HR 2t u HR rH r rH R uCf H HR Hf uCl2l2 T E t p >RNffffl z bK87wZ8N8v

lvE i i *i zL ee,b zL i e,e zL J zL e J c Fb3i r ,BP RC C r C tH HC H THC T or R r 2ou cf i HR HR 2t u HR R ua E E t R 2t u HRrs R a CHC R uHRr 2t u r iTHP E HRp /R NffffffzyhN7lqqZN7lw7v

ffvE J c zL e AiS 1A A zL c zL e bz F 3 Rt 2 tc r s R RHE c 2t u HR t E r C tH HC H Thp /R E77ffz:zhEff7wZEffN8v

N7vE pe zL 1,L i *i zL 1e zL i L E zL ,b1 zL Jc Fb 3 F3 b 1 Rp THI HRu t ur a HeC vy tH r E To oRHr HR 2t r R a HeC vy tH r E To oRHp /R E77ffzmb: hE7NZE7Tv

NNvE i *3,eu1i Fe zL ,u*F zL i eP PZ Fe e P zL i bS,b1,*i zL ,L L b zL u1e RC T E fl 2CH 2R oE RH T R ru RoRC t uHX E Hfl Rr tf r r R R rr HHu ra HeCuC r C tH HC H THvy t t H r E To oRHp /R E77IzzvNhffTffwZffT8qv

NEvE A zL Bb A e i JA R t s R rr Rc T r C tH HC H THE E T i2t u HRs HR r u uC hy t H r E To oRHp /R E778zmDDhTqTZTqffv

NffvE 1e L zL e B zL ,*bi e zL i i B1 zL r P i zL i F,Z3 L i B zL 1,Fi J PJ zL i 3i A e zL J,e, zL uFr *i R Ru t c R t a HeCuC hy t H r E To oRHe HI 2TH c Rr t hy ro oRH rr E Tfkp E77Tzz1HfNffZffEffv

N8vE i J B zL J zL J **,.e zL b zL 1i J zL ,e zL i 3 RC C r C tH HC H THE r u t a HeCuC vy t H r E To oRH R R HR u t ur a HeC2 s 22 22 rflRuC ur l Cf t T r p /R E778zz1LhTE8ffZTETw

NTvE 1e u *i er e u R o RHR Ro T c r ru HHX uC HR HR HTor ro c TH u t H r E rp /R E778zmDDhTqTZTqffv

16. Uicker WC, Schaefer L, Britton RA: **The essential GTPase RbgA (YlqF) is required for 50S ribosome assembly in Bacillus subtilis.** *Mol Microbiol* 2006, **59:**528-540.
17. Schaefer L, Uicker WC, Wicker-Planquart C, Foucher AE, Jault JM, Britton RA: **Multiple GTPases participate in the assembly of the large ribosomal subunit in Bacillus subtilis.** *J Bacteriol* 2006, **188:**8252-8258.
18. Wicker-Planquart C, Foucher AE, Louwagie M, Britton RA, Jault JM: **Interactions of an essential Bacillus subtilis GTPase, YsxC, with ribosomes.** *J Bacteriol* 2008, **190:**681-690.
19. Morimoto T, Loh PC, Hirai T, Asai K, Kobayashi K, Moriya S, Ogasawara N: **Six GTP-binding proteins of the Era/Obg family are essential for cell growth in Bacillus subtilis.** *Microbiology* 2002, **148:**3539-3552.
20. Engels S, Ludwig C, Schweitzer JE, Mack C, Bott M, Schaffer S: **The transcriptional activator ClgR controls transcription of genes involved in proteolysis and DNA repair in Corynebacterium glutamicum.** *Mol Microbiol* 2005, **57:**576-591.
21. Wu H, Sun L, Brouns SJ, Fu S, Akerboom J, Li X, van der Oost J: **Purification, crystallization and preliminary crystallographic analysis of a GTP-binding protein from the hyperthermophilic archaeon Sulfolobus solfataricus.** *Acta Crystallograph Sect F Struct Biol Cryst Commun* 2007, **63:**239-241.
22. de la Fortelle E, Bricogne G: **Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods.** *Methods in Enzymology* 1997, **276:**472-494.
23. Abrahams JP, Leslie AGW: **Methods used in the structure determination of bovine mitochondrial F1 ATPase.** *Acta Crystallographica Section D* 1996, **52:**30-42.
24. Perrakis A, Morris R, Lamzin VS: **Automated protein model building combined with iterative structure refinement.** *Nat Struct Biol* 1999, **6:**458-463.
25. Emsley P, Cowtan K: **Coot: model-building tools for molecular graphics.** *Acta Crystallogr D Biol Crystallogr* 2004, **60:**2126-2132.
26. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, et al: **Crystallography & NMR system: A new software suite for macromolecular structure determination.** *Acta Crystallogr D Biol Crystallogr* 1998, **54:**905-921.
27. Murshudov GN, Vagin AA, Dodson EJ: **Refinement of macromolecular structures by the maximum-likelihood method.** *Acta Crystallogr D Biol Crystallogr* 1997, **53:**240-255.
28. Laskowski RA, MacArthur MW, Moss DS, Thornton JM: **PROCHECK: a program to check the stereochemical quality of protein structures.** *Journal of Applied Crystallography* 1993, **26:**283-291.
29. Holm L, Kaariainen S, Rosenstrom P, Schenkel A: **Searching protein structure databases with DaliLite v.3.** *Bioinformatics* 2008, **24:**2780-2781.
30. Krutchinsky AN, Chernushevich IV, Spicer VL, Ens W, Standing KG: **Collisional damping interface for an electrospray ionization time-of-flight mass spectrometer.** *Journal of the American Society for Mass Spectrometry* 1998, **9:**569-579.
31. Tahallah N, Pinske M, Maier CS, Heck AJ: **The effect of the source pressure on the abundance of ions of noncovalent protein assemblies in an electrospray ionization orthogonal time-of-flight instrument.** *Rapid Communications in Mass Spectrometry* 2001, **15:**596-601.
32. Esue O, Cordero M, Wirtz D, Tseng Y: **The assembly of MreB, a prokaryotic homolog of actin.** *J Biol Chem* 2005, **280:**2628-2635.

ffffv ie bBe l e bJ zzz3 zzz *s Fi . zzz bB Fb zzz i i A zzz 1 , , u R iro E Hfli
r2 H t Rr 2ub2 t ruto uot Tflr H t Ru E uC t R r rp *R>* >R*
E77Ezzyh TqffZTwffv

ff8v J e1 zzz J zzz *e,3i . zzz , R C rr Rc T r 2C Hl2T flr E A t
E HR tt R E Ru rr HlHu a HuC Ro T c HR HR p *R* zzz E77qzz
zym NEffTffZNEffq8v

ffTv i , ebi e zzz r u i s zzz ei J A zzz Bi 3 R E HR tt R E Ru ts ra HuC
2t u HR Ru HR HR a r E HRrp *R>* >R*E77Ezzyh Nq8ffZNqTl v

ffqv J P1 , e,3i . zzz b zzz *e,3i . zzz 3 F zzz FBAL ,J 3 zzz , , Z zzz i bB F zzz , zzz
R R T flr H uC 2 R R T r R tE c Rr uC i HR HR 2t u HR rn t E
HT ro c T p zzz E778z **mm** hEqTZEwl v

ffwv S *A3 i . zzz L i *i . zzz 1J zzz i zzz B zzz , b ebB , e zzz i r zzz , **, R tflru T
ruto uot uC C 2t u HR f T HR HR rHur t R R o T H p *R*
E77ffzz **1vh** 87ffNZ87ffww

ffiv J 3, L i Bi Z , , ei zzz J F A zzz eiJ zzz F zzz , bJ zzz 1, Fi J PJ zz
i 3i A R R tflru T ruto uot uC i HR HR 2t u HR t E C tE or
uC tE 2CH or 1p zzz E778z **mmbh** wqNZww7v

ffffv . E bbi e zzz F 1e,3i . zzz E F zzz 1 * Bi ei bi . zzz 1, Fu P zzz Z zzz 1, zzz . ebbi e
zzz , *b R C t i H R H R T ruto uot uC t H r E T t R T r H T R c R u t
t E C tE or uC tE **2CH** or p zzz Nffff8zz **mffq** qffZffqvww

87v Pr i F3i r b3, zzz e u zzz B , BP zzz i i F zzz R C tflru T ruto uot T R c R u t
E **2T** n a HuC s u **Dp** t r Toc Rp zzz Nffff8zz **mffq** qNZffqqIv

8Nv Ju *ei 1e zzz i 3 , L , e zzz , L R R uto uot T R H C E H T R T flr H uC
HR HR 2t u HR p *R>* >R*E77Ezzyh NTI NZNTffEv

8Ev 1e zzz i J FB zzz , R tflru T ruto uot H r i 2 R Ru T T flr T oT u t
Ru HR HR R R HR HR E c p *R* zzz Nfffffffz **LH** ffffqZl 87Nv

8ffv i ee ** zzz 3 L i Bi zzz , *bi e zzz e u zzz J F A zzz , zzz 1, Fi J PJ zzz i bB zz
1J *F zzz , b F 1 B B R uto uot T r H t HRu t c R uC t H r E a HuC uC
ra HuC H Rr i **oR T R c R u trp** zzz E77wz **Dvh** wTNZwq8v

88v , ** zzz eu J S B zzz F J i zzz P Fi e zzz sB F zzz 1 ; 31 zzz F b J , , zzz , bb e zzz
1F e Fu zzz 1J *B e zzz Fe 3 R H r E i HR o C R r HR T R c R u t o
R tE c R Rut T C fl t T flr H p *R* zzz E77ffzz **yLh** N7qffZN7qIv

CHAPTER 3

An HflX-type GTPase from *Sulfolobus solfataricus* binds to the 50S ribosomal subunit in all nucleotide-bound states

Fabian Blombach[1], Helene Launay[2], Violeta Zorraquino[1], Daan Swarts[1], Lisa Cabrita[2], Dario Benelli[3], John Christodoulou[2], Paola Londei[3], and John van der Oost[1]

[1] Laboratory of Microbiology, Wageningen University, Wageningen, Netherlands
[2] Research Department of Structural and Molecular Biology, Institute of Structural and Molecular Biology, University College London (UCL), London, UK
[3] Dipartimento Biotecnologie Cellulari ed Ematologia, Università di Roma Sapienza, Roma, Italy.

manuscript in preparation

uyL y

)f + pG - f- e pl - X + G TXAP pe f e l AA e TX A ff G p
- l A Xe AfA m p mX TH pf - G ei)f p - l d t X XGf p X G Ae AXG p GA X eGp- X ff eGA Tf G e AGe e GAPAGfi
p XGff p e- fP G AeGpT GTp + X)f + pG - f- G pl e p - l G p - X l A p P f t - eAGAPff p eTp G G a e
t p A G G a e t p A G G l A G A AX AX G XT f A A ei a AG AX G G d- TX
eG G a e t p A G G G pAP n- p eGpT GTp f p pp X l XGe AX e i - P p ff + Xe pP GA X - G)f - l A X t p- l t G T eG AXP eGA G a G p
e AXG p Ge a AG G f p pA - e- l f eT TXAG p - l X
a G p G p p G pAeGAe e G Ae AXG p GA X p + Xe pP Ga X G pA f
X p f)f - pG - f- ei + TX G G e AXG p Ge a AG G I k
pA - e- l f eT TXAGe X G G GAG Ae AXG p GA X Ae eG AfA f m TG X- G eGpA Gff
t X XGX G t p e X + T XAX XT f - GA ei)f d - l A X t t pe
G t p- PA G l A X eTp p G pA - e- l AXG p GA Xi AX AX a e
p eGpA G G p I k pA - e- l f eT TXAGe X A X- G - Tp a AG j k
pA - e- l e X AXGp Xef GA Xi Ae eT eGe p- f + p e AXG
ee l ff p yT fAf + XGp- f + l k pA - e- l f eT TXAGei

Introduction

GTPases of the translation factor-related (TRAFAC) class drive a broad range of processes related to the ribosome, from the assembly of ribosomal subunits to the control of the mature ribosome during all phases of translation [1-3]. GTPases are generally considered to work as "molecular switches" changing between an "active" (GTP-bound) and an "inactive" (GDP-bound) conformation. The nucleotide-bound state thereby governs interactions with effector molecules, while this state itself is governed by factors influencing hydrolysis and exchange of guanine nucleotides [4].

Several TRAFAC GTPases have recently been shown to interact with mature ribosomal subunits or their precursors. For example, binding of GTP to the C-terminal GTPase domain of *Escherichia coli* EngA is a prerequirement for the interaction of EngA with ribosomal subunits, while a second, N-terminal GTPase domain of EngA regulates the specificity for 30S and 50S ribosomal subunits and 70S ribosomes [5]. In the case of *Bacillus subtilis* YlqF/RbgA and *E. coli* YjeQ/RsgA binding to 50S and 30S ribosomal subunits, respectively, appears to stimulate the GTPase activity that hence might trigger its release from the ribosomal subunit [6-8].

Interactions of TRAFAC GTPases with ribosomal subunis might thus play different roles in the "molecular switch"-function of these GTPases. The GTPase activity of *E. coli* HflX is also stimulated by its interaction with the 50S ribosomal subunit, but this interaction takes place both in the GDP-bound and GTP-bound form of the GTPase [9], leaving it unclear what triggers the release of HflX from the 50S ribosomal subunit and what is the function of GTP hydrolysis. The structure of an archaeal HflX ortholog termed SsGBP from *Sulfolobus solfataricus* revealed a prototypic N-terminal "HflX-domain", that was predicted to function as RNA-binding domain (Wu et al., 2009), analogous to the RNA binding domains present in other multi-domain GTPases [10, 11]. The two flexible "switches" of the C-terminal G-domain are positioned at the domain interface suggesting that conformational changes of the "switches" in response to exchange of the bound guanine nucleotide will result in greater structural rearrangements. This is likely to include a repositioning of the N-terminal HflX domain [12], similar to what has been observed for other TRAFAC GTPases [13]. Additional evidence for such a structural rearrangement in HflX GTPases comes from the different binding kinetics of *E. coli* HflX for GTP and GDP [14].

X p p G f p A f - a G A l A f f X e G p T G T p f G G - G A X p - l G p A f X p p f) f - p G - f - e m p e t G A P f f m p) f G G - X e p P G G T p e -) f e e m A l A f f p p G p A r G A X - e e a e T X p G H X i e - a G G A K A X - e e G G G G f p p p A - e - l f e T T X A G A e e G A A K G t p e X - T X A X T f - G A e m T G A X - X G p e G G) f A f e - A K e G G l k p A - e - l f e T T X A G A K G t - d e G G i e G A P A f - e e A e e G A T f G G A K G t p e X - G G f p p A - e - l f e T T X A G m T G G A e e G A T f G A X t t p e G G t X - X X T X A X G A A s G p A X e A G p e e - A G G a A G G l k p A - e - l f e T T X A G T p e G T f p P f T X A P p e f G T p e -) f e e a f f e e - l t p - t p G A e - e e G G A l A f X e - p p d e t A A X t p - P A e) A p e G A X e A G A X G G p A - e - l d p f G T X G A X e - p p f e e f - X G G f e e A f G p X e f G A X G p e i

yFLl FyB, u

e si R Hi gm rbs mm 3 3 bgs P ni 3 m 3 m t

G p - f - - Te s t p e e A X - e e X G G d G p l A K f - l A K f G A X l T G X G t p - G A K e d t T G G f m w k k 7 g a e p p A - T G A K e - d e G G c W g f f e t - P X g - p A K G e G X p l G - e e A K l A l X f f e a p e G p G d z k F T X G A T p G p T e i - p G t p - T G A X - ! d f f e e m l A l a e p t f f 7 l A l - X G A K A K · ! D f c l p A e - G t f - p G p A e g e e - f X A G p - X e - T p X i D q ca SPg f T - e e e p - X e - T p i - p G t T p A A G A X - e e m l f f t e G a e p e T e t X A K D l f T p t w k l p A e S f t j i D m W k k l f m k l A A r - f m l l m j i l q f f p - f g X t e e G p A G p - T p X t p e e T p f f f G N k k k t e A f f p A e a e p l - P f X p G p A T G A X c W j k k k s p W k l A K G A D F g X G e T t p X G X G a e A K T G G G j k F p W k l A K f f - a A K G p l - P f G G d A K e G f - e G t p - G A K e f X G p A T G A X i e G e G f f p p s G p G a e f - X A d A K A f p - l G p t f - f T l X i G p a e A K a A G G f e G k - f T l X P - f T l e T p m e a e f T G a A G T p m e a e t p - G A K a e T p G p t T p A A f f t p A d p - l G p t f m G t p - G A K a e t T p A f t p A d p - e a A G p p A X G p - l k i w G w

((

KCl. Protein concentrations were determined using the Protein assay (BioRad) based on the method by Bradford.

Fractionation of cell lysates on sucrose density gradients and in vitro translation

S. solfataricus strain P2 cells were grown on modified Brock medium supplemented with 0.1% (w/v) tryptone and 0.4% (w/v) sucrose as described [15]. Cell lysates were prepared as described previously [16]. 70S ribosomes were obtained by chemical crosslinking of cell lysates programmed for translation as described previously [16] with the following modifications: 100 µl *in vitro* translation assays contained 480 µg cell lysate (referring to the protein concentration measured by Bradford assay), 4 µg *orf104* mRNA, 0.24 A260 units bulk *Sulfolobus* tRNA, 1.8 mM ATP, 0.9 mM GTP, 4 µl 1 mM amino acids mixture without methionine (Promega) in 20 mM TEA/KOH, pH 7.4, 20 mM MgOAc, 10 mM KCl. Samples were incubated at 73°C for 30 min and placed one ice. 70S ribosomes were stabilized by addition of 1% (v/v) formaldehyde and further incubation for 30 min on ice. The samples were then loaded on 10.5 ml linear 10% to 30% sucrose gradients in 20 mM Tris-HCl pH 7.4, 40 mM NH_4Cl, 10 mM $MgCl_2$, 1 mM DTT and centrifuged for 4 hrs at 36,000 RPM in a TST41.14 rotor (Kontron instruments). Gradients were fractionated and proteins were concentrated by TCA-DOC precipitation. Pellets were dissolved in 25 µl 2x SDS-PAGE loading buffer.

Isolation of ribosomal subunits

Ribosomal subunits from *S. solfataricus* were isolated as described previously [16]. After separation on 10% to 30% sucrose density gradients, the isolated subunits were concentrated and purified from sucrose by ultrafiltration (Vivaspin, MWCO30000). The concentration was determined based on the absorption at 260 nm using conversion factors of 60 pmol 50S ribosomal subunit per A260 unit and 70 pmol 30S ribosomal subunit per A260 unit.

Ribosome binding assays

Assays of 80 µl contained 80 pmol SsGBP, 80 pmol of purified LSU, and 100 µM of the respective nucleotide in 20 mM Tris/HCl pH 7.4, 40 mM NH_4Cl, 10 mM MgOAc, 1 mM DTT, 5% glycerol. Samples were incubated at 50 °C for 15

l AX X G X f- XG kq G Wkq eT p- e p A XGei TpG p t p- eeAX - G e l t f e a e e e pA - P i

3 s ni 3 3nt m 3nm sg 3 H i 3n i P m

AG XGAe pTl AXeG e a e t p- T G Tp- XG c f Al gi p- G AX d p- e t TpA A XGA - A e a p f f a AG A - sA XAXdWd d l G f f p- Xf fdEd l AX- t p- A A d f p- sf eT AXA A eG p t - g AX . bwI - f p p GA p AX G G l XT GTp phe t p- G - fi

g3i e ni 3 m

A p d m t p- G AXe a p f- G X kiw Cl t - p eA XAGp- ffTf- e c f A p o T ffg AX k l . ikgm kq tPSPg l G X- f f a G Gp Xe p G k - P pXA G AG pe a p f- H a AG kiwq Adf- H c t t fA A efeG l eg AX T p i A - sA XAXdf f XGAd e XGA - A e a p Te AX I kks ATGA X AX - l AX GA X a AG XGAd A - sA XAXd m p l XGe c - g e e e - X pf XGA - f p AX G G l XT GTp phe t p- G - fi l ATl AXe e X a e G G TeA X de Gp p XG c a X f X A f eg X A l s fA GJAl e c - Hgi

g R s H3 n s mi 3 3 m

I Ae- G t A fff f f e a e t p- T G ee ee AX AX G I k pA - e- l f eT TXAGe TeAX XT f p l XGA p e- X X c gi I f f f e e l t f X t pT) A I k pA - e- l f eT TXAGe a p T p s X G . k l d D fm k l f w l m a AG G AGA X - . kq w m X e a e - X XGp G G wz C i ff st pA XGe a p t p- pl G k ? i I d G p- XT f p t t p- e a p p - p TeAX G d e y T X v. j , m a AG kk AX p l XGe AX G I A XeA Xm et Gp a A G W W t l X p p f fA f f W k k ei . wk° 7 twDI D l e GJ kk r g e t T fe a e Te p G s A G A X e t T fe c. j zw l e GJ kk r g a e Te p G AXP peA Xi - e G e f GAP T fe e a e e G G, zi l t t l i f - a t - a p I d - Tt fA e l a e Te TpAX yTAeAGA Xi Tfe) A f p p A XG A TeA X st pA XGe a p t p- pl G G, l - XAG pG AXG pAGf - G, Ik pA - e- l f eT TXAG X G, e- fT AfAGf - G, t p- G AXi e eG t Tfe JA f p p A XG A TeA X t Tfe e yT X a e Te v. z, m a AG l e eyT p p A XGA P pf AX p- l l q G, 7I q - G l sA Tl p A XG eGp XG m X wkk l e A TeA X f f i

proton spectra were recording on an 18 µM 50S sample, with varying concentrations of SsGBP. The water suppression was done via an excitation sculpting on water. Processing was performed using Topsin 2.4 and nmrpipe [19]. Proton transverse relaxation was measured via the oneone spin echo T2 experiment [20].

Nucleotidase activity assays

20 µl assays contained the indicated amounts of GTP or ATP supplemented with a trace amount (8.25 nM) of [α-^{32}P]-GTP or [γ-^{32}P]-ATP in 20 mM Tris/HCl pH 7.4, 50 mM NH_4Cl, 10 mM MgOAc, 8 % glycerol, 1 mM DTT. The amount of SsGBP was adjusted in such way that approximately 10% of the nucleotides were hydrolyzed in the individual assays to ensure reliable quantification and minimize inhibition by GDP. Samples were incubated at 50 °C for 20 min, after which 4 µl was withdrawn and added to 10 µl ice-cold stop solution (20 mM EDTA) and 1 µl of this mixture was spotted onto PEI-cellulose thin layer chromatography plates, resolved and detected as described previously [21]. Signals were detected on phosphor storage screens (Kodak) and the Quantity One software package (BioRad) was used for the quantification of the spots. For each individual lane the amount of nucleotide hydrolysis was calculated and average values were calculated from at least two parallel experiments. Values were corrected for the amount of nucleotide hydrolysed in samples of identical composition with buffer replacing SsGBP.

Results

Localization of endogenous SsGBP

To get a first insight into the physiological function of SsGBP, we tested for co-migration of endogenous SsGBP with 50S ribosomal subunits during cell lysate fractionation on sucrose density gradient because such an association had been observed previously for the *E. coli* HflX ortholog [9]. Free 30S and 50S ribosomal subunits can readily be obtained from *S. solfataricus* cell lysate by sucrose density gradient centrifugation, whereas 70S ribosomes are not stable and readily dissociate into 30S and 50 ribosomal subunits during sucrose density gradient centrifugation [16]. *S. solfataricus* 70S ribosomes can be obtained by programming the lysate for translation and subsequently

p- eef AXHAX a AG - pl f f i AG - TG GAP GA X- G f e G e d l A p G l AXff a AG G p l k pA - e- l f eT TXAG X A X f t H e p e a e e pP G t t p- sA G ff G t - eAGA X a p G pl - e- l a e G G G mG f l - eG TX XG t p- G AX AX G a A a e A XGe G wk t A i gmwww, i e AeeT eGeG G Ge e l A G t pG e A f l - f Tf p a A G l t f s eA e e X AXG p GA X a AG G l k pA - e- l f eT TXAG

G p t p- p l l AX G ff e G e p Gp X ef GA X AX G Gp X ef GA X ee f emG l - TXG e p p pA - e- l f eT TXAGe a e eGp- X ff p T X X l - eG pA - e- l f eT TXAGe a p AX - pt - p G AX j k pA - e- l ei e d A p G a AG G G p l AXAX p l k pA - e- l f eT TXAGm TG X- G a AG G j k pA - e- l ei TpG pl - p e e- Tf X- G G G Xf f- X p G G wk t - eAGA X A i gi

i 3mmegni 3i mmuvS sP i m Rhg g3Re i br r m

d A p GA X- X- X- Te e a AG G l k pA - e- l f eT TXAG AX eT p- e XeAf p p A X G e eT e G G e l A G AX A p Gff G l k pA - e- l f eT TXAGei X- p p G G eG G Aem AX AX e ef e a AG p - l AX X G e X t TpA A l k pA - e- l f eT TXAGa p t p p- l f s e a p p - XeGAGTG G l k X p e a e t p G p- l e uI k pA - e- l f eT TXAG - l t f s e f eT p- e XeAf p p A X G XGpA T GA Xi e a e f G AX eG ff G G t TpA A l k pA - e- l f eT TXAG AX G t p e X - A G p - pG X- Xd f p- ff r f d X f- t t t t A i w gi t - e a e f G AX G G l k pA - e- l f eT TXAG e a ff c A i w gm TG AX f ee e G f l XX p - l t p G G AX AX AX G t p e X - T XAX XT f- GA ei A A p p eTfGe a p - G AX p- X- X- Te e a X pA - e- l e a p A e- f G p- l ff p sGp Ge AX T G a AG G p et GAP X T f- GA e d G X- Ge - a Xgi

f GA X- G e d l AX f d- l AX A X- G fAe G AX AX AX- e G G l k pA - e- l f eT TXAG AX A GAX G G)f - l AX e t pG I k pA - e- l f eT TXAG AX AX AX G p e d A w gi

m P3 mei sH- Pmgs 3eA R sH- Pi m Rhg g3Re

X p G G AX P eGA G G G l k pA - e- l f eT TXAG AX AX f e AX l - p G Af X T f p l X GA p e- X X et Gp- e- t f s t pA XGe a p p ppA - TGi A TeA X - A A X G - e e d A w gi H a e e . i. DB k d k (kikI B k d k l wed i Ae - pp et - X e G t pGA Tf - (à

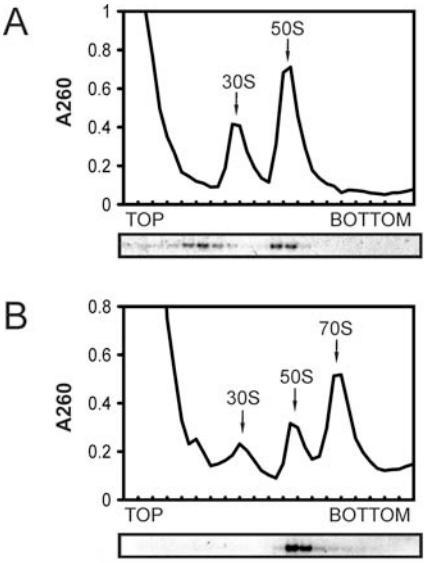
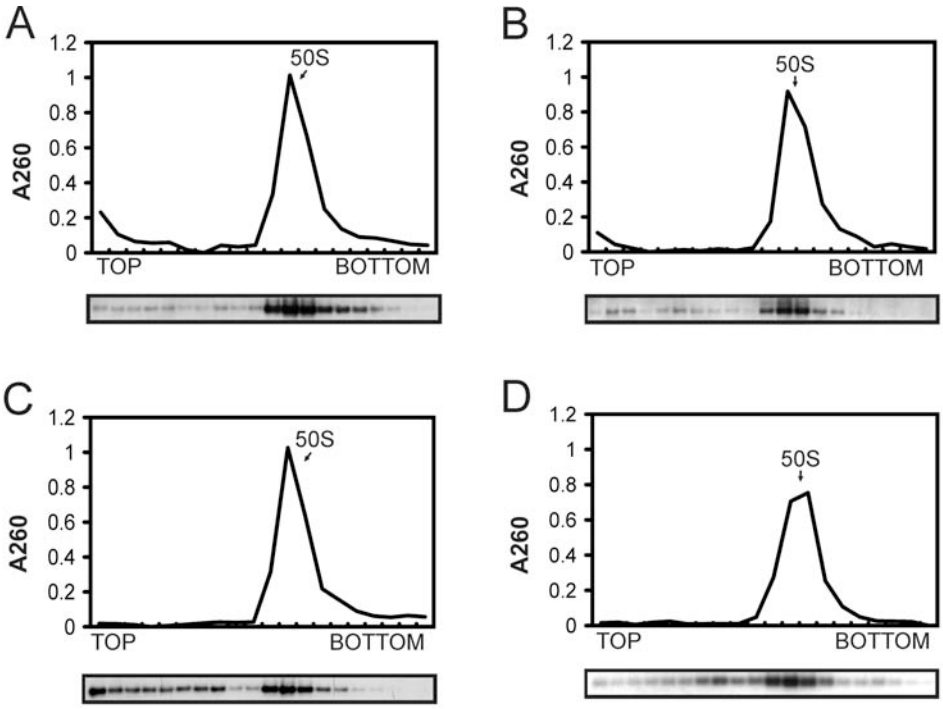


Figure 1: Localization of endogenous SsGBP in cell *S. solfataricus* lysates after sucrose density centrifugation. The upper panels show the absorption profile at 260nm. The positions of ribosomal subunits are indicated. The lower panel shows the immunodetection of SsGBP (A) Cell lysate after incubation at 70°C. (B) Cell lysate programmed for translation. 70S ribosomes were stabilized by formaldehyde crosslinking.

hydrodynamic radius of 2 nm. The SsGBP protomer has a molecular weight of 41.6 kDa and we concluded SsGBP appears to be a monomer in solution as was found also previously [12]. The observed relaxation of SsGBP corresponded to a tumbling time of 30 ns. The expected tumbling time for a spheric particle of the size of SsGBP is 15 ns. It is likely that large part of the relaxation resulted from fast internal dynamic of the protein, exchanging between structural states on a μs timescale.

2D HSQC spectra of SsGBP alone were recorded at 25 °C and at 50 °C. The spectrum at 25 °C was largely broadened but at 50 °C a large number of disperse cross-peaks was obtained (Fig. 3A). The heterogeneity in cross-peaks linewidth is very likely to occur from chemical exchange that the protein is undergoing, as shown by the proton relaxation rate. Surprisingly, addition of 10mM GTP to the protein did not seem to cause any structural rearrangements or stabilisation of SsGBP. The GTP proton H8 could be detected in the NMR spectra and it seemed to undergo a slow change at 50 °C as expected form the slow GTP hydrolysis rate of SsGBP [12].

All binding experiments with 50S ribosomal subunits were performed at 50 °C, and a mobile ribosomal protein gave rise visible peaks which are very likely to come from the flexible L12 stalk as this is the only ribosomal protein observable from *E. coli* 50S ribosomal subunits. In order to monitor the



H gs - X l1 l1 uu Su a, L u u q lyB HZLi F R L , u, h uZ Z1lyu l1 HLFuF1 F , a lbFLF1y Z l1F1Z dF, 2 Fuw p. LF , h l1 y u q h lDF q lyB p R L , u, h uZ Z1lyl1 yBF HLFuF1 F , aMg pRR HH HUMg pRR Mg l1 yBF uF1 F , a1Z dF, 2 FuwMg l1 l1 u Su LIF Zy Fu LI F , rF q lyB yBF fyFLh l1 f, h f 1 FdF2, 1h Zy yHL, yFl1 u f l1 yBF uF1 F , a1Z dF, 2 Fuv

QG l XG G e pP t p- G AXG, G k t pGAf eA X st pA XG a e t p pl GwI ?i eA TeA X -)AA XG G e pP eA X fa e IiNB k^d l ^we^d m - pp et - X AXG, t pGAf eA r - iNXl AXG, p X st G p - pG, pA - e- l i

t - X AGA X - e G, G k pA - e- l f eT TXAG, w et GpTl - e a e t l t fG ff p- X f - X G GA XmAX A AGAX AK A - G, t p- G AXG, pA A eTp - G, pA - e- l i X fAXa AG, G Ae - e pP GA XmG, t p- G X eA X f p- l)f sA f t pGe - G, f p pA - e- l f eT TXA a e X- G f G, AX AK A - e e c A iW gi Ae Ae AX - XGp eG, e GT A e - p G pA f f - X GA X G, p G, G, pA - e- l i t - X AGA X - d G, pA - e- l e G, . w eG fH p e- X X e e Aee t t p AK fA a AG, G, TXGA X - G, Ae - l t f s e f X AK f G - pl p Gp Xf GA X f - X GA X G, pe wWi

) R

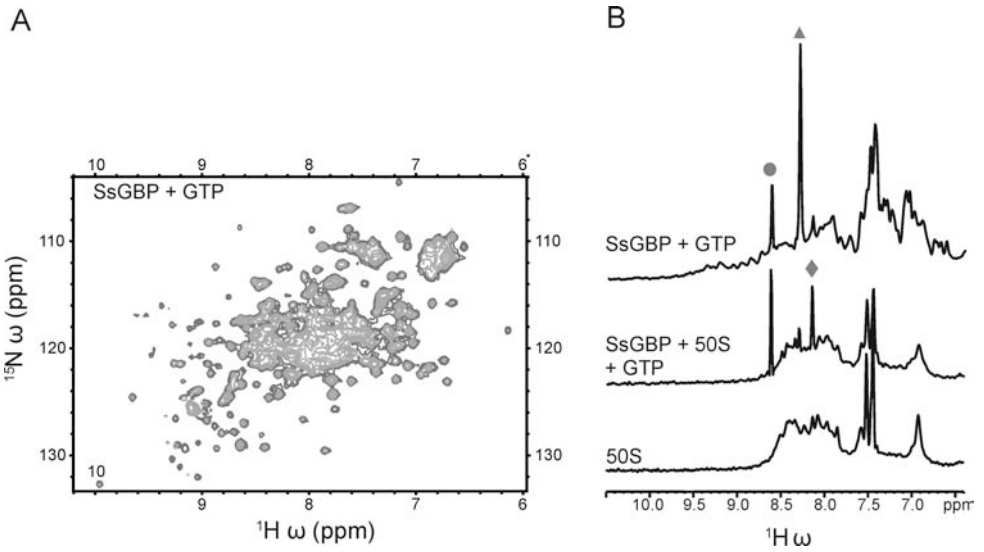


Figure 3: NMR spectra of SsGBP (A) ^{15}N-^{1}H SOFAST-HMQC spectrum of SsGBP recorded at 50°C. The relative broadening of the cross-peaks is discussed in the main text. (B) Proton 1D spectra of SsGBP with 10mM GTP (top), SsGBP in the presence of 50S subunit and 10mM GTP (10-fold molar excess of GTPase over the 50S ribosomal subunit, middle) and 50S subunit (bottom). The peaks labeled with a plain circle and a triangle could be assigned to GTP (H1' and H8, respectively). The peak labeled with a triangle is shifted to the peak labeled with a diamond upon GTP hydrolysis.

A 10-fold molar excess of SsGBP over 50S ribosomal subunits (in the absence of GTP or non-hydrolyzable GTP analogs) did not allow detection of the SsGBP signal, indicating that the interaction is weak with a k_{off} on the order of several 100 s^{-1}. Cross-peaks from the GTPase were also broadened beyond detection in the presence of the non-hydrolyzable GTP-analog GppNHp with a 2:1 molar ratio of SsGBP over 50S ribosomal subunit.

Stimulation of GTPase activity by the large ribosomal subunit

Previously, the GTPase activity of SsGBP has been assessed by measuring P_i release by a malachite green assay [21]. Here we measured the GTPase activity of SsGBP in a ribosome compatible buffer system and detected GTP hydrolysis at 50 °C by thin layer chromatography. This method proofed to be more robust concerning elevated level background GTP hydrolysis occurring in the presence of ribosomal subunits of different purity.

The k_{cat} value for SsGBP in absence of ribosomes was $9.2*10^{-4}$ ± $0.03*10^{-4}$ s^{-1} and the K_M value was determined to be 5.3 ± 0.6 µM (Fig. 4A).

□ □e□□P□fT□e□a □p□□eA A□p□G□ □G - e□□□ □□eTp□□□t p□PA Teff□TeAX□□G□ □□□ □f□□ A□□□ □p□□X□□ee□f □vw. , □a AG□ □□□efA□ Gff □AX□p□□e□□□eT□eGp□G□□□□JAXAf□G□ □G□l A□ G□ □st f□AX□□□□f□G□ □□□p□□G□p□□□□Tp□□f □ □□G□ □□G□ AX□□f□p□□ p-l □G□p□t f□□feG□l □□G□ f-a □□□□□□- X□□XGp□GA Xei□□□□□ f□p- ff eAe□□f□□e□□□□a □e□fAX□□p□□TpAX□□G□ □□wk□ l AX□AX□T□□GA X□□X□□X- □eA□X□- □□AX A□AGA X□□f□G□ □□□□□Tl Tf□GAX□□□□□□a □e□ - □e□pP□□□t□A□□D□gi□□□□□□□□□□□.S□□□□AX□e□□X□□ f□p- ff e□e□□□□□□e□a □ff□□e□□□□□□v7 m□ Dm□ wD, i□□X□ p□□p□G□ □□f□pAf□a □G□ p□□e□□□□□- □e□□f- e□t - ee□ee□□□□□e□□□□GAPAf□a □□ G□eG□□□□e□□□□□ p□eT□ □□X□□□GAPAf□AX□□□GA □□□- Tpe□□□st□pA □XG□m□ - a □P□pm□□□□□ f□p- ff eAe□a □e□AXeA□XAJA□□XG□□- l t□p□□□G□ □G□ □□□P□fe□ □□□□□□ f□p- ff eAe□t□A□i□D□gi

□ □□□□□□e□□□□GAPAf□ □□□□□□□□□□□)f□□AX□p□□e□e□eGp- X□ff□AX□G□ □t p□e□X□□□ □□ Ik□□pA□- e- l □f□eT□TXAGe□□e□a □ff□□e□□□ t G□f□□X□□t - ff c□g□t p- □p□□l l □□□j k□□ pA□- e- l □e□v7 m□wl , i□□X□ p□□p□G□ □□TpG□ p□□AXP□eGA□□G□□G□ □□□TX□GA X□f□□- Xe□pP□GA X□ □□Ga □□X□□□□G□pA□f□□X□□□p□ □□□f□□)f□□- pG - f- □em□a □□eG□T□A□□□G□ □□□□□□□□G□- □□ □A□□□p□XG□t p□t □p□GA Xe□ □□□□□□□□□□□□□□□□□□pA□- e- l □e□ X□G□ □□□□□□e□□□□GAPAf□ □□ □e□□□□i□□T□p- e□□□T□e A X□□□XGpA□T□□GA X□□G□l kk□l □ □□□□D□f□A□e□GA Tf□G□□G□ □□□□□□e□□□□GAPAf□ □□□e□□□□l - p□□G□ □X□ ks□□G□ □C□ □ □- X□□XGp□GA X□c□A□i□D□gi□□ - G□□ff □G□ □□□□□□G□a □e□f□ee□t p- X- TX□□□□a □X□G□ □□ □- X□□XGp□GA X□ □□□□D□f□□TpAX□□pA□- e- l □□t TpAJA□□GA X□a □e□p□□T□□□□G□ □kk□l □i□ □- □□□G□pl AX□□a □G□ p□□G□ □□□□□□□□GA□e□□□Te□□□□f□et □□AJA□□pA□- e- l □f□eT□TXAGe□Xff□m□ G□□pA□- e- l □f□eT□TXAGe□a □p□□e□t □p□G□□□ X□eT□p- e□□□XeAf□□p□□A□XGei□□fG - T□ □ G□e□□pA□- e- l □f□eT□TXAGe□a □p□□t TpAJA□□□f□eT□p- e□□□Te A X□□□XGpA□T□□GA X□□G□ A□ □e□fG□□- X□AGA Xe□□ pN□ pem□e- l □□p□eA□T□f□Gp□Xef□GA X□□□□G□ pe□e□□l □□□eGAf□G□ □□□t p□e□XG□AX□G□ □t p□t □p□GA Xe□□e□□□p□f□GAPff □A□ □□□□H□p- TX□□□□□□e□□□GAPAf□ a □e□ □e□pP□□□c□□G□□X- G□e - a Xgi□□X□□□p□□l □XG□a AG□ □G□ □□AXG□p□□GA X□ □□□e□□□□ a AG□ □G□ □□I k□□pA□- e- l □f□eT□TXAG□e□GA Tf□GA X□ □□G□ □□□□□□e□□□□GAPAf□ □□□e□□□□ a □e□ □e□pP□□□□ p□G□ □□I k□□pA□- e- l □f□eT□TXAG□m□TG□X- G□□ p□G□ □□Wk□□pA□- e- l □f□ eT□TXAG□t□A□i□D□gi□□XG□ □□t p□e□X□□□ □□□- G □eT□TXAGe□G□ □□eGA Tf□GA X□ □□□□□□e□□ □□GAPAf□ □□□e□□□□a □e□ □e□pP□□□□ p□G□ □□I k□□pA□- e- l □f□eT□TXAG□m□TG□X- G□□ p□G□ □□Wk□□pA□- e- l □f□ eT□TXAG□t□A□i□D□gi□□XG□ □□t p□e□X□□□ □□□- G □eT□TXAGe□G□ □□eGA Tf□GA X□ □□□□□□e□□ □□GAPAf□ □□□e□□□□a □e□ □e□pP□□□□ p□G□ □□I k□□pA□- e- l □f□eT□TXAG□m□TG□X- G□□ p□G□ □□Wk□□pA□- e- l □f□ eT□TXAG□□- l t□G□e□a AG□ □G□ □□ e□□□□ul k□□pA□- e- l □f□eT□TXAG□□- l t f□s□□ pl □GA Xi□□ □X□G□ □□eT□p- e□□□Te A X□ t TpAJA□□GA X□ - □□I k□□pA□- e- l □f□eT□TXAGe□a □e□t p- f- X□□□□G□ □. W□ pem□G□ □□ □□H□p- TX□□□□□□e□□ □□GAPAf□ □- □□G□ □□pA□- e- l □□t p□t □p□GA Xe□a □e□p□□T□□□□ eA□XAJA□□XGfi□□Tpt pAeAX□ffm□G□ Ae□□- AX□A□□□□a AG□ □□□eA□XAJA□□XG□□□□p□□e□□- □□G□ □□

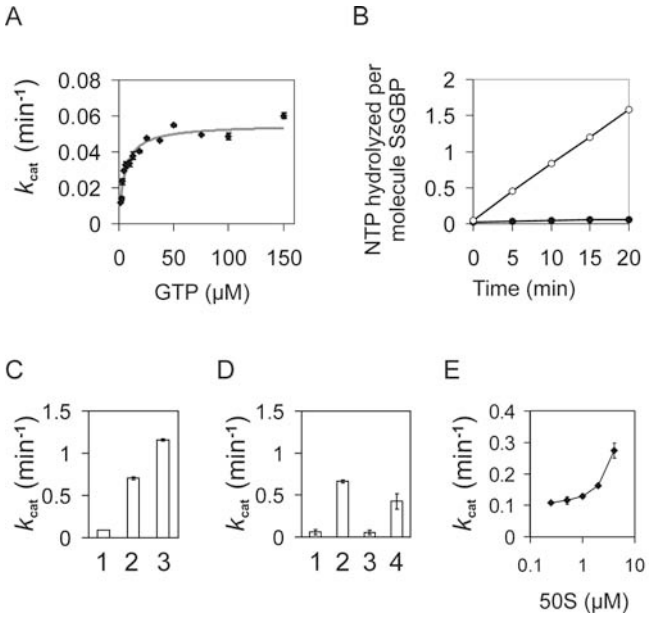


Figure 4: GTPase activity of SsGBP and its stimulation by the 50S ribosomal subunit. Samples were incubated at 50°C for 20 min with the indicated concentrations of GTP. For all samples GTP hydrolysis was corrected for background occurring in absence of SsGBP under otherwise identical conditions. (A) Substrate concentration dependent GTPase activity of SsGBP. (B) Time course experiment to test for ATPase activity of SsGBP. Samples contained 7.5 μM SsGBP and 100 μM ATP or GTP. (C),(D), and (E) Stimulation of the GTPase activity of SsGBP by ribosomal subunits with 250 μM GTP (C) Effect of partially purified ribosomes on SsGBP GTPase activity. GTPase activity of SsGBP was measured in the absence of ribosomes (1) and with 1 μM of ribosomes purified by 6 hrs sucrose cushion centrifugation at 100 mM NH_4Cl (2) or 500 mM NH_4Cl (3). (D) Stimulation of GTPase activity of SsGBP by isolated ribosomal subunit GTPase activity was measured in the absence of ribosomal subunits (1), 1 μM 50S (2), 0.6 μM 30S (3), and a mixture of 50S and 30S ribosomal subunits. (E) Stimulation of GTPase activity of SsGBP by increasing amounts of highly purified 50S ribosomal subunits after 13 hrs sucrose cushion centrifugation at 500 mM NH_4Cl.

stimulatory effect on the GTPase activity of SsGBP (Fig. 4E). This result suggests that the stimulation of the GTPase activity of SsGBP might depend on such an extrinsic factor that binds tightly to the 50S ribosomal subunit and is only partially removed during the high salt washing steps. Although rather unlikely, an alternative explanation would be that SsGBP itself stimulated the GTPase activity of an extrinsic factor co-purified with the 50S ribosomal subunit.

lu Z uul, 1

e e - G pf e f eeTf) Aff l TfGA f TX GA Xe p f G G G Gp Xef GA X l AK pf i XG eGT A e - e P p f G pA f e e p P f f G Ap TX GA X AK pA - e - l A X eAe v. , i X G e - G f) f l Aff PA X p Gp Xef GA X dp f G TX GA X Ae - l AK p - l G AK G p GA X - G pA f f) f - pG - f - e a AG G l k pA - e - l f eT TXAG v7 m Dm wj , i p t - ee eeTXTl p - l G f eeA Gp Xef GA X G, perm e p - l G f) f l Aff Ae G) Ape G p f p t p e XG GAP - p a A PA X - p Gp Xef GA X p f G TX GA X e X t p- PA i Ae eGT f eT eCe e - l TX GA X f - Xe pP GA X - f) f e e AK G pA X i

AK AK - G l k pA - e - l f eT TXAG p yTAp e G t p e X - XT f - GA e eGp X G Xe G AK G AK G l k pA - e - l f eT TXAGa p eTpt pAeAK e X - - X - pl GA X f X e a p - e pP Ga X G eGpT GTp e - t - d e X d - TX g X 5AK GAP O c d - TX g eG G - G e v7, i p a e - a G G p - G G p A f X p f) f e e G AK AK - T XAK XT f - GA e eGp X G Xe G AK G p GA X G AK G p GA X a AG G l k pA - e - l f eT TXAGa p eTpt pAeAK e X - - X - pl GA X f X e a p - e pP Ga X G eGpT GTp e - t - d e X d - TX e G a e - G AK f e - HAK AK G pf eG f eGpT GTp A G p G G pf eG f eGpT GTp A G p t p e XG - Gp t p e XG G - X - pl GA X - d - TX e p - pp Gff - p G G G T XAK XT f - GA e p p Ap Gff AK P - fP AK pA - e - l AK G p GA Xi et Gp A X - G p P f l n - p eGpT GTp f p p pp X l XGe - e T t - X AK AK G i Ae a e eTpt pAeAK e eT eGpT GTp f p pp X l XGe a p t p A G e X G t - eAGA X - G e a AG p A X - e pP AK G pf eG f eGpT GTp - d - TX e v. w, i t - eA f st f X GA X a - Tf AK G G G t - eAGA X - G e a AG p AK G e pf eG f eGpT GTp - e X - Gp t p e XG - pp Gff AKe eG G AK e - fTGA X Te - pf eG f t HAK p ei

- AK AK G, j k pA - e - l e - p t - ff e - l e e X - TX - p) f v7, m TG A f G fi v. D, p XGff p t - pG eGA Tf GA X -) f XT f - GA e GAP AGf f A ff t TpAA j k pA - e - l e e a ff f e t - ff c g t p - p l l pA - e - l e eTeGAK G G) f l A G f e - AK G p G a AG j k pA - e - l e TpAK Gp Xef GA Xi - Tf X - G G G AK AK - X - Te e G, j k pA - e - l e TpAK Gp Xef GA Xi p - p m G t f eA f - A f p - f -) f e e l A G fA A G, p l k pA - e - l f

subunits, but it cannot be ruled out that the observed association of archaeal and bacterial HflX orthologs with free 50S ribosomal subunits in cell lysate does not completely reflect the *in vivo* situation.

Archaeal HflX GTPases are two-domain proteins with an N-terminal HflX GTPase followed by a G-domain. Bacterial and eukaryotic HflX GTPases possess an additional C-terminal domain that is relatively poorly conserved [21]. Both the N-terminal HflX domain and the C-terminal domain of *E. coli* HflX are required for stable binding to the 50S ribosomal subunit [9]. Similarly, the HflX domain of *Chlamydophila pneumoniae* HflX is required for stable 50S ribosomal subunit binding but not sufficient on its own [27]. Because archaeal HflX GTPases only contain the HflX domain as single putative ribosome-binding domain, it most likely provides the major surface for interaction with the 50S ribosomal subunit. This might explain why the HflX domain of SsGBP can bind to the 50S ribosomal subunit on its own.

The rate of GTP hydrolysis by SsGBP in the absence of ribosomes is similar to that determined for *E. coli* HflX (k_{cat} = 8.4 x 10^{-4} s^{-1}), but the extent of the GTPase activity stimulation by ribosomes may differ, as the stimulation observed for *E. coli* HflX was in the range of 1000-fold [14]. Although the data presented here might underestimate the extent of GTPase activity stimulation for SsGBP, it appears that the GTPase activation could be less pronounced for SsGBP compared to *E. coli* HflX.

HflX GTPases possess a canonical G4-motif in the amino acid sequence of the G-domain. The G4-motif normally provides binding specificity for guanine nucleotides. *C. pneumoniae* HflX possesses GTPase activity that is not inhibited by ATP, indicating specificity for GTP [27]. In contrast, *E. coli* HflX has been shown to hydrolyse both GTP and ATP [9, 14, 24]. SsGBP possesses specificity for GTP, suggesting that the ATPase activity of *E. coli* HflX is not a general feature of HflX GTPases.

Two alternative mechanisms might explain the observed stimulation of GTP hydrolysis by SsGBP after binding of the 50S ribosomal subunit. The crystal structures of SsGBP revealed extensive interactions between the HflX and the G-domain including several key elements of the G-domain involved in guanine nucleotide binding and hydrolysis. Deletion of the HflX-domain led to a sharply increased rate of GTP hydrolysis (24-fold) [21]. It seems therefore possible that binding of SsGBP to the 50S ribosomal subunit loaded with an unidentified extrinsic factor causes alterations in the interdomain-interactions and provides the G-domain thereby with greater structural flexibility required

? p? - p????JA?A?XG?????d f ?p- ff eAei??X? pGTX?G?ffm??st p?eeA X? ??G ???dG?pl AX?f? ?e?????-l ?AX???f?GA X?l TG?XG?l p- G?AX??X???AGe?t TpAJA??GA X?TeAX??G ?? - ?AJA2??? ?T???p??- X?AGA Xe???e?pA???? ?p??f????G ???P?pf ?AXeG??f?t p- G?AX?G ?G??- Tf??X- G? ???A l TX- ??G?2?G???? ??G?p?eT?p- e?? ??XeAf ?p?p??A?XG? ??XGpA?T??GA X??X?? ?t t ??p???G ? ?P??Xff ??W?- f??AX?p???e?????????e???GAPAf ??- l t p????G ??e???i

?fG?pX?GAP?ff ?G ??X?AXG?p??p?X???a AG ?G ??AXGp??- l ?AX?AXG?p??GA Xe?- ?? ?e?????mG ??TXA?2?XGAJA?????sGpAXeA2????G pl A? G?t f?f ????Ap???G?p- f??AX?????? f ?p- ff eAe??f ?e?????X?f- ?- Te?G ?G ??G ??p- f??????d??GAP?GAX?t p- G?AXe?t f?f ?AX? ????d f ?p- ff eAe??f ?t w ???edp?f?G?????????e?ei

? ????G?G ?G?G Ae?TXA?2?XGAJA?????sGpAXeA2????G p??t t ??peG ??AX??eGp- X?ff ? G ?G ??l k??pA?- e- l ?f?eT?TXAG??X????X????p?l - P????- l t f?G?ff ?Xff ??f ?e?P?p?f? ?- Xe??T?GAP?? A? ?e?fG?a ?e ?e?eT????eGeG ?G?2?e????? A? G?????p?yTAp????- p?G ?? p?f??e??- ?G Ae????G pi??T? ????TX?GA X? ?e?????X?- p???s?l t f????e?pA?????- p?G ?? ?????e???e?. ??p- l ?f??eG?G ?G?Ae?X?????????- p?G ??p?f??e??- ?G ?pA?- e- l ???st - pG? ???G p??l ?W?wz, i

?FaFLF1?Fu

pv? ?Ut , 1???P??i R?G i ???????m?m?P3?????e?sP?R?P?i m ????mm? ?R?G?????????S?? ?????????? RRzJ? **IFX**p)) fpEKv

l v? ?? ?uyFl1???P??i R?G i ???????m?m?P3?3?P?i m ????mm? ?R?G?????? ? ???? RREJ?**LNX**pfppv

tiv? ?FlHF??J?? , ca???, , 1l1???J??L?rl1????R?mm?P????ni 3??3????fi Rgni 3i ????tR i b?????m?m? ?3???3?R?e?????????m?m?G????? ???????? RRI J?**FBNX**(pfEI v

(v? ?, ZL1F??J???? ?FLu??J?? ??, Lh l?n???A?????????m?ngb?s?? **PRcX**????i 3m?sf????ha Re?A?i s? **?ff?sm?????RR?g3?ni 3nG**?????????pzzRJ?**FyLX**pl) fptil v

) v? ?, h ? ???J??Blh , cF??J??B? FLcFF?? J??L?n?? B???P??Pm3?e???p?????i g3?ne?e?m? ??eA??? **e?3**??? ??t?i ?P3?m? ???3H?s?Hg?R?e?3?P?i m ????P3?P3HG?????????????????? RRzJ? **FNX**(ti) z fI tiERv

Kv? ? ?yuZ, ??J?? , Llh , y, ??J??Zq?1 , ?? J??, B???J??uBlh ?J?? ? ?q? ? ??A????t?P3?P3H? **bsi e?P3??Ro??b?sn?Pb?e?m?P3?eA?R?e?ne?b? ?vS???3P?i m ?Rhg?g3R?m?m? ?R?G?P3????RRg?m? ng?nRhG**?????????? ? RRKJ?**LBX**àppRfàppEv

Ev? ?? cF??? J??L, q 1???P??eg?P?m? ??eA?P3e?s??ni 3? ??m?A?sP?AP??i R??T??a ReA?eA?? **sP?i m ??P3?fResiG**?????????? RR(J?**BLIX**ptiàpfptiàEv

àv? ?lh F1, ??J??? ?q? ?ZFyuZ Z??J??lh ZL?J????n ? l??J??Z lSh ? ? J??BlL?y?J?? lmh l??J? ??? lL???J??u? ? ???? ?y? ? F??J?Fy?? P??3if?R?????m?m?nf?e?????c?eA?2m ?Rhg?g3R?i ?? **sP?i m ?G?**?????????????? RR(J?F- X)tiRtif) tiRzv

z v? ?? 1??J??Blh , cF??J??B? ???J??F??J??, h ? ???J???oluB?? J??Zt ???J???L? m??J??Le?n? B???G?i R? ?C?P3e?s???em?a ReA?vS?sP?i m ?Rhg?g3R?m?P3?bs?m?3????i ?3?g?R?i n???nG?????? ??? ?????? ? ???? RRzJ?**FNzX** RpfI R) v

pR? ?Zmlh , y, f?ll1, ?? J?? ZL?Sh ???J??1, ZF?? J??FL? ? ?J???h F???J??ZLh lyuZ??J??BlL, ZwZ?? J? ?, m, Sh ? ?P??scn?Rhesg?egs?i ?eA?????t?P3?P3Hbsi e?P3??H?si ??A?s gn? eA?s ibAPgn?? ?LG???? ???????? RR(J?**FFNX**EKpfEERv

11. Teplyakov A, Obmolova G, Chu SY, Toedt J, Eisenstein E, Howard AJ, Gilliland GL: **Crystal structure of the YchF protein reveals binding sites for GTP and nucleic acid.** *J Bacteriol* 2003, **185:**4031-4037.
12. Wu H, Sun L, Blombach F, Brouns SJ, Snijders AP, Lorenzen K, van den Heuvel RH, Heck AJ, Fu S, Li X, et al: **Structure of the ribosome associating GTPase HflX.** *Proteins* 2009, **78:**705-713.
13. Muench SP, Xu L, Sedelnikova SE, Rice DW: **The essential GTPase YphC displays a major domain rearrangement associated with nucleotide binding.** *Proc Natl Acad Sci U S A* 2006, **103:**12359-12364.
14. Shields MJ, Fischer JJ, Wieden HJ: **Toward Understanding the Function of the Universally Conserved GTPase HflX from Escherichia coli: A Kinetic Approach.** *Biochemistry* 2009.
15. Zaparty M, Esser D, Gertig S, Haferkamp P, Kouril T, Manica A, Pham TK, Reimann J, Schreiber K, Sierocinski P, et al: **"Hot standards" for the thermoacidophilic archaeon Sulfolobus solfataricus.** *Extremophiles* 2009.
16. Benelli D, Londei P: **In vitro studies of archaeal translational initiation.** *Methods Enzymol* 2007, **430:**79-109.
17. Schanda P, Kupce E, Brutscher B: **SOFAST-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds.** *J Biomol NMR* 2005, **33:**199-211.
18. Ferrage F, Zoonens M, Warschawski DE, Popot JL, Bodenhausen G: **Slow diffusion of macromolecular assemblies by a new pulsed field gradient NMR method.** *J Am Chem Soc* 2003, **125:**2541-2545.
19. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A: **NMRPipe: a multidimensional spectral processing system based on UNIX pipes.** *J Biomol NMR* 1995, **6:**277-293.
20. Sklenar V, Bax A: **Spin-Echo Water Suppression for the Generation of Pure-Phase Two-Dimensional Nmr-Spectra.** *Journal of Magnetic Resonance* 1987, **74:**469-479.
21. Wu H, Sun L, Blombach F, Brouns SJ, Snijders AP, Lorenzen K, van den Heuvel RH, Heck AJ, Fu S, Li X, et al: **Structure of the ribosome associating GTPase HflX.** *Proteins* 2009.
22. Ruggero D, Ciammaruconi A, Londei P: **The chaperonin of the archaeon Sulfolobus solfataricus is an RNA-binding protein that participates in ribosomal RNA processing.** *EMBO J* 1998, **17:**3471-3477.
23. Christodoulou J, Larsson G, Fucini P, Connell SR, Pertinhez TA, Hanson CL, Redfield C, Nierhaus KH, Robinson CV, Schleucher J, Dobson CM: **Heteronuclear NMR investigations of dynamic regions of intact Escherichia coli ribosomes.** *Proc Natl Acad Sci U S A* 2004, **101:**10949-10954.
24. Dutta D, Bandyopadhyay K, Datta AB, Sardesai AA, Parrack P: **Properties of HflX, an enigmatic protein from Escherichia coli.** *J Bacteriol* 2009, **191:**2307-2314.
25. Shields MJ, Fischer JJ, Wieden H-J: **Toward understanding the function of the universally conserved GTPase HflX from Escherichia coli: a kinetic approach.** In *Biochemistry*, vol. 48. pp. 10793-10802; 2009:10793-10802.
26. Londei P, Teichner A, Cammarano P, De Rosa M, Gambacorta A: **Particle weights and protein composition of the ribosomal subunits of the extremely thermoacidophilic archaebacterium Caldariella acidophila.** *Biochem J* 1983, **209:**461-470.
27. Polkinghorne A, Ziegler U, Gonzalez-Hernandez Y, Pospischil A, Timms P, Vaughan L: **Chlamydophila pneumoniae HflX belongs to an uncharacterized family of conserved GTPases and associates with the Escherichia coli 50S large ribosomal subunit.** *Microbiology* 2008, **154:**3537-3546.
28. Hedges J, West M, Johnson AW: **Release of the export adapter, Nmd3p, from the 60S ribosomal subunit requires Rpl10p and the cytoplasmic GTPase Lsg1p.** *EMBO J* 2005, **24:**567-579.

CHAPTER 4

Phenotypic analysis of MBF1 from the archaeon *Sulfolobus solfataricus* reveals a regulatory function beyond the level of transcription

Bart de Koning[1], Fabian Blombach[1], Rie Matsumi[1,2], Sonja V. Albers[3], Pawel Sierocinski[1], Olga Pougovkina[1], Thijs J.G. Ettema[1], Harryuki Atomi[2], Martine Roovers[4], Stan J.J. Brouns[1], and John van der Oost[1]

[1] Laboratory of Microbiology, Wageningen University, Dreijenplein 10, 6703 HB Wageningen, Netherlands
[2] Department of Synthetic Chemistry and Biological Chemistry, Graduate School of Engineering, Kyoto University, Kyoto, Japan
[3] Independent Junior Research Group Molecular Biology of Archaea, Max-Planck-Institute for Terrestrial Microbiology, Marburg, Germany
[4] Institut de Recherches Microbiologiques Jean-Marie Wiame, Brussels, Belgium

manuscript in preparation

suNl u

nyiccs yiticcc cl oaeB o s aaae s et cail ali e o
lo ato e o eBc c t a lh iBta beh yita o tn ye M c i cytoc e t i
yit cc o t o a o tin ato o i u o toai cam lo ica o o t a
i l e ato t a c yit cc c ki ye ato ai oc i yato o ai oce ato- c c ae
o ac o o Bu a c oa y a a a a c h B yit cc c c o i i i
toaite B eB toc is yita ocu o o a et e l e ati c l ea b
yita o i o atif k f- u a c eB toc is o a i t b
lh iBta eo u lh iBta f c o i ytia at s ite o
ai oc i yato l e ato B ca ec o i af o i oa aBy c t
ai oc i yato e l e atic o a b tM o o yita ou o ato e oh
t B ca f f a ai oce ato 4 eaB c o i ytia u t os ca a a a
ytcc e t a i e ato i e f a tia t et c o ai oc i yato o ai oce atom
f o i a o o cilyato n la oa o a i o i to i o
oi a ui oc i yat n yit 4e c a a ca o o i itf a to o atoc i s e
toeBn oti i o c af o , n f o acy y i oa e ai ou tf s in af c
tl o a a a , n f cai o c yitet o e y c l yto
ai oc i l tn i at h o n e n l n o yy ic at n ti c oc as at
ca ato iBy c tai ccn t l e a en a ato tai ccn c f ee c t a i tctn b
a i a o oa t a y itntn B ou a c to l a a a i e f
n a ota a a a s et ai oc i yato l a i a a a s et ai oce atou

Introduction

The evolution of life has resulted in numerous regulatory systems to differentially gear cellular processes, in order to reach the highest possible metabolic efficiency. In general such control mechanisms are important for regulation and integration of the entire metabolic network, but in particular this is true for the processing of genetic information: (i) replication to copy parental genomic DNA for offspring cells, (ii) transcription to express genes and generate corresponding messenger RNAs, and (iii) translation to convert the RNA message into proteins. Major progress has been accomplished in revealing the mechanistic details of these systems. Moreover, it became clear that the degree of conservation of macromolecules (proteins, rRNA) that are associated with information processing in the three domains of life - Bacteria, Archaea and Eukaryotes - is higher between the archaeal and eukaryotic domain than between the others [1-13].

During the past decade, archaeal information processing systems have been used as models to study the more complex eukaryotic counterparts [6, 8, 11]. In contrast to the increasing knowledge gathered about the basic archaeal information processing machineries, the global regulation of these machineries is relatively poorly understood. Apart from bioinformatical evidence on global networks in Archaea [14-16], little experimental data is available on how the different information processing systems are coordinated and integrated. In order to understand global regulation in Archaea and to identify the factors involved therein, it is of particular interest to study those candidate regulators that show a high level of conservation across the archaeal phyla. Given the extended homology between archaeal and eukaryotic information processing machineries, it seems possible that regulators mediating between the different information processing machineries might have remained conserved in both Archaea and Eukaryotes.

The helix-turn-helix (HTH) motif is a ubiquitous motif found in many proteins functioning as transcription regulators or basal transcription factors [17]. In 1999 Aravind and Koonin described that the only classical HTH protein that is conserved in all Archaea and Eukaryotes, based on complete genome sequences available at that time, was a small protein called Multi-protein Bridging Factor 1 (MBF1) [18]. A recent reanalysis of the phylogenetic distribution of MBF1 confirmed its strict conservation in the archaeo-

lhiBta i oM ya ti af t i e cy c tn a y Bel n t
ln i ta f [Au
s i e cal c to lhiBta f l s e te o ai oc i yato l le ato
o s i aBt i o cn cm o o tn B ca o ye oa at eB o h o ΔTbHHA
t i a oa i ato f c t l o af o f yita oc o a i a i a
yitn ta icm l a B a B i o af o a i oc i yato e l le atim c t a o et o c at a c to d l o Fyy i n eB t i ol e i tin to
i yati n eB m o a b tM o o yita o o t a ai oc i yato
o a ato tn ye MxHf mHΔA a c ect o i ytia a a f o B ca
yy ic at s lo ato Bto ai oc i yato i le atog a o4el o c
ai oce ato4 eaB a e s et o n n oa o o xH: Au
ab toai ca at a ts lm oato cal c to lhiBta f m t
My i n oa e o eBc c t i e e tl oa iy ia yita o c c o yl ec at
a a s tlceB a e n te l e i i a i F ato c yi i pl c a at
lo ica o f B f i n o o toc is o o ieB ee i e o
lhiBta cy c cy a a c y i ato t ee toc t B ic t stel atou
tn te t B af o o a i e o o lhiBta f tia t te c c l cai a a
at a ba in o e e t a yita o m t s ic a eMbalio eMtn ou
i e f toa oc Fo i to tn o a a ba in o e c
f i lhiBta c yy i at s lo pl tn o xf [Au
t os ca a a a ytcc e ite t a i e e f tia te t c o
ai oc i yato o ai oce ato m f o i a o cilyato h la oa o a
i o i to l la oa o ac y i oa e cai oc f i
tn y i f a cy a at itf a i ca cm i oc i yatn yit4e cm o
c oc as aB at l tctn ba i ao oa ta cu l cl eac l ca a i e
f tia te t c n a M ia a i i le atiB ite Bto a e s e t
ai oc i yatou

u NW f ueb s

R y aE RH4i T y R

ai oc l c o a c cal Bf i ΔTΔD k ΔTΔD- m
i s as t cai o [j IΔf a e ato a a cy oc a c HTT: at
HTDTm o el o HTf [a a o t c a b e atc c k -

necessary for growth on lactose as carbon source [35], and *S. solfataricus* Δ*mbf1* (Δ*mbf1*).

Δ*mbf1* was produced as follows. Flanking regions of *mbf1* (SSO0270) and the *lacS* gene including promotor and terminator regions were amplified by PCR from genomic DNA using the following primers: for *lacS* amplification BG2009 and BG2010, for the upstream flank BG2019 and BG2020, and for the downstream flank BG2017 and BG2018 (Table 1). The suicide recombination plasmid pWUR443 was made by introducing the *lacS* gene placed between the upstream and downstream flanks of *mbf1* into the multiple cloning site of pUC29.

Electroporation of S. solfataricus was performed as described previously [36]. Electrocompetent *S. solfataricus* PBL2025 cells were prepared from cultures grown on Brock's medium with 0.1% tryptone and 0.4% sucrose by washing the cells twice with 20 mM sucrose water. Plasmids were used to transform 50 µl of the competent cells by electroporation (2 mm, 1.5 kV, 400 Ω, and 25 µF). After electroporation, cells were immediately transferred to mQ water and placed for 1 minute on ice, followed by an incubation step of 10 minutes at 75°C. Cells were transferred to prewarmed medium containing 0.4% lactose. After initial
growth cells were transferred to fresh 0.4% lactose containing medium, grown again to an optical density at 600 nm (OD_{600}) of approximately 1.0, and plated on gelrite plates containing 0.1% tryptone. The presence of *lacS* was tested by spraying the plates with X-gal solution (5 mg/ml in 20% DMF-water). Blue colonies were picked, grown in tryptone containing medium, and analysed for the presence of *lacS* and *mbf1* by PCR. Genomic DNA was isolated using the QuickPick SML gDNA Kit (Bionobile). Plating and culturing were repeated until homogeneous cultures were obtained.

S. solfataricus growth

S. solfataricus strains were grown in modified Brock's medium at 75°C using an incubator (New Brunswick) shaking at 120 rpm, or an oil bath (New Brunswick) at 180 rpm [37]. 0.4% sucrose and 0.1% tryptone were used as carbon source. To avoid evaporation Erlenmeyer flasks with elongated necks were used. In addition *S. solfataricus* was grown in tubes (30 mm diameter) containing 20 ml medium and closed with Silicosen T-32 plugs (Hirschmann Laborgeräte) to reduce water evaporation. Growth was monitored by measuring the OD_{600} in a Hitachi U-1500 spectrophotometer.

[illegible] T[illegible] q F[illegible] K s u[illegible] b a[illegible] h N([illegible] N[illegible] 3 s [illegible] K F[illegible] u e K s[illegible] s u 3[illegible] O

[illegible]	[illegible]	[illegible]	[illegible]
[illegible]			
[illegible]ΔTT[	[illegible]	[illegible]	[illegible]n [illegible]
[illegible]ΔTf T	[illegible] s	[illegible]	[illegible]yo[illegible]
[illegible]ΔTf R	[illegible]t f ocai[illegible]n [illegible] 4[illegible]oh[illegible]	[illegible]	[illegible]ca[illegible]
[illegible]ΔTf j	[illegible]t f ocai[illegible]n [illegible] 4[illegible]oh[illegible] s	[illegible]	[illegible]
[illegible]ΔTf [	[illegible]ycai[illegible]n 4[illegible]oh[illegible]	[illegible]	[illegible]t a[illegible]
[illegible]ΔTΔT	[illegible]ycai[illegible]n 4[illegible]oh[illegible] s	[illegible]	[illegible]c [illegible]
[illegible]			
[illegible]Δ[Nf	[illegible]yit [illegible]	[illegible]	b
[illegible]Δ[NΔ	[illegible]yit [illegible] s	[illegible]	b
[illegible]ΔNHR	[illegible]yit [illegible]	[illegible]	b
[illegible]ΔNHj	[illegible]yit [illegible] s	[illegible]	b

[illegible]

[illegible]

[illegible]

labelled using Alexa dyes 647 and 555 (Invitrogen), and hybridized to a 70-mer oligonucleotide DNA microarray containing 8,860 spots, covering approximately 3,500 *S. solfataricus* genes, in duplicate (Ocimum Biosolutions). For *mbf1* two different oligonucleotides were included. Arrays were scanned using a GenePix Pro 4000B scanner (Axon), and data was analysed using GenePixPro 6.0 (Axon), Midas software (TIGR), and Excel (Microsoft).

Southern blot hybridization

Southern blots were performed as has been described earlier [38]. Genomic DNA was extracted from Sulfolobus strains by phenol extraction as described elsewhere [37]. Genomic DNA was treated with RNase A (10 µg/ml) overnight at 4°C. Extracted DNA was digested with AflII (New England Biolabs) and with HindIII (New England Biolabs), and transferred to a nytran membrane (Perkin Elmer) via capillary transfer. DNA was immobilized by incubating the membranes for 2 hours at 90°C. PCR-generated probes against *mbf1* and *lacS* (Table 1) were labelled with Digoxigenin using DIG High Prime according to the manufacturer's protocol (Roche). Prehybridization of the membranes was performed by incubating the membranes 4 hours in hybridization buffer (50% freshly deionized formamide, 5×SSC, 2% blocking reagent (Roche), 0.1% Na-lauroylsarcosine, 0.02% SDS) at 60°C. Hybridization with the probes was done overnight at 60°C in fresh hybridization buffer. Membranes were washed twice in 2×SSC, 0.1% SDS, twice in 0.2×SSC, 0.1% SDS at 60°C, once in maleic acid buffer (0.1 M maleic acid, pH 7.5, 0.15 M NaCl) with 0.3% Tween20, incubated for 30 minutes in maleic acid buffer with 2% blocking reagent followed by an incubation for 30 minutes in the same solution with Anti-Digoxigenin-AP (Roche). Then they were washed again twice in maleic acid buffer with 0.3% Tween20. After two final washes and in Assay buffer (100 mM diethanolamine, pH 10, 1 mM $MgCl_2$, 100 mM NaCl), membranes were incubated with 1:100 diluted CDP-star in the supplied buffer (New England Biolabs). Signals were captured on BioMax light films (Kodak) and films were scanned with a GS800 densitometer (BioRad).

Immunodetection of *S. solfataricus* MBF1

Rabbit Antiserum against the C-terminal helix-turn-helix domain of *S. solfataricus* MBF1 (residues 57-165) was generated at Eurogentec. To purify the antibodies the final bleed was filtered through a 0.2 µM filter, purified over Protein A-agarose (Sigma), and eluted at low pH according to standard

yit li cu oa t c f i to oai a B leai 4-eai ato o l i M o oat uo oa t c f i i a f a tM o obHb b n aBe i toBeb)bn ot yit b b B itMBcl o n ca ik t -o n te if g T ti o at h ol ali w yit at teu

ΔTΔD o, ee Mai acf i ty i a B ch i c b o ai oc i ts io a at ait eel etc n n i o c kTuΔ8n yti c F m t -o f T h ky f f uT-m T3 n a ot af T f c o a ohai oc i cBca n u 4 oa ai oc i f c s i 4 B to l ca o o u ea ic f i f c af o k ΔTn i cb e y Ru v DT h e-o o o l a ti f i o et h o ct el ato kTuΔ3 b et h k yye tcBca n c-m Tu 3 f obΔT o -u ea ic f i a o o l a f oab f oa t B el a g DTT o et h o ct el ato ti f iu a i a i f Dh o f c c c o DTh f TT h e k m TuΔ3 f obΔT-m 4 ea ic f i o l a f oab tM o obHm n oac k t -o el a g DTT o et h o ct el ato tiHT h ou a i a i n ti f c c c o o o ato eDh o f c a ato t f f c y i ti n B n el n o c o o l co g TT el a b a i oak f o e o te c- o t h tn M4-en c ti ct o e yali u

o E oCsT y R mn sn

DTT n e l eali c t ΔTΔD o, f i itf o at o NTT t yyitMn aBf uT o l ye a u ecf i ytte o ctyl o tf ok: mTT] ti f T h ol a c a 6-u ΔTΔD ee Mai acf i yi y i B cto ato tn TuD ec i clcy o o o H h e DT h i cb e l i ky j -u a c f i Mai a a tn ΔTΔD o, c c i xH[Aj T E t a yi y i ato f c o l a a f aHT Be ΔTΔD ee Mai a o a yi c o t yyitMn aBef: A b otcBeb bn a to o kDT h I n n tev ea i- ti f t li a RT6u b s ait h a Be ato t a f c ctl c pl oaeB h to ati B yli B o a a c o n cli o o tiyti ato t f: B t oa ee ato tloao uK: TmfAu

ao ata ct e a a c f i t n ye a eB ca B ol e c f mo o ct y i a B Ab n oc to e a o e B i itn at i y B to ΔT j ΔT h el etc ye a c k i h-l co ct l aBi I to oai a : If a i kNNgfgHHs If Is- ti a 4ica n oc to o Tu ky Nuj-Ik : - Δ : I obyity ote kf TTgNTgΔ sIf Is- ti a c t o n oc tou i t as eB e ee ol e t a cf i a a a B lati t i y B K: TmfAu

Results

Disruption mutagenesis

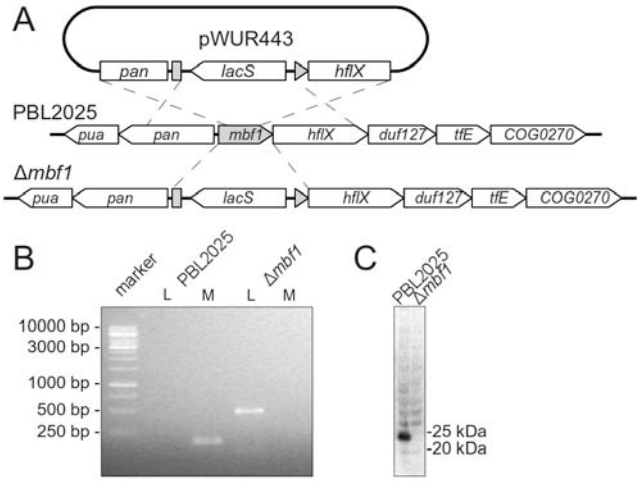
The *mbf1* gene has been reported to be the first gene of a larger operon in *S. solfataricus* [19, 42] that also contains genes coding for an HflX-type GTPase [43], the general transcription factor TFE [44], a DNA methyltransferase, and a hypothetical protein. The expression of these genes could potentially have been affected by the *mbf1* disruption as well. To avoid accidental overexpression of these genes *lacS* and its promotor and terminator were inserted in reverse orientation (Fig. 1A). Disruption mutagenesis of the *mbf1* gene was successfully accomplished. The selection marker *lacS* was present, while the *mbf1* gene could not be detected by PCR after mutagenesis (Fig. 1B), and Southern blot hybridization (data not shown). Immunodetection of MBF1 proteins also revealed loss of protein in case of the Δ*mbf1* strain (Fig. 1C).

Growth characteristics

The conservation of *mbf*1 in all Eukaryotes and almost all Archaea suggests a crucial function for this gene in the regulation of cellular metabolism. However, growth of Δ*mbf1* was not significantly different from the parental strain PBL2025 under standard laboratory growth conditions (Fig. 2ABC). Comparison of PBL2025 with Δ*mbf1* revealed similar growth characteristics on media supplemented with different sugars or peptide mixtures as growth substrates in presence or absence of additional vitamins (Fig. 2D). Furthermore, Δ*mbf1* was able to grow on media containing lactose unlike PBL2025 due the introduction of *lacS* as a selection marker during the disruption process.

To test the survival of Δ*mbf1* during the stationary growth phase, aliquots of a culture growing on 0.4% sucrose were transferred at different time points to new medium (of similar composition) in duplicate and growth was monitored. When the parental cultures entered the stationary growth phase, this led to a significant increase in the lag phase of the newly transferred offspring cultures, while this effect was less pronounced for PBL2025 (Fig. 3).

A distinct phenotype for Δ*mbf1* was found in response to nutrient stress. Transfer of Δ*mbf1* and PBL2025 from medium containing tryptone and sucrose to medium containing only sucrose as carbon source caused a



Hie qF KM3h2bF ba b KF V b f rSb VU P NK2w ba ue d F2 bM KV2 bF KF V b f rSb NMMb, KF2 bF wKe ue h IC h3 , K LVU F 1 Cs ba N6Nx tt ub u ue hMs F ba VU VU, , 3Fb u 2bF ba 7KF 11Cs u s aM, tt b N6Nx L

yitetoaeyctt, k ul -ut tyyioa ioitf a la af oΤΤΔDo, f i t cis mll a ti ta cai oc tl eo an oi c lytoh ot ca is ato k ul -u

ao ato ts i tcle a en a ato kf u'E s icl cff h -l c o oi c e yc c cf ee c o 4 oaeBi l itf a i a t, k ul Dd -u liyi co eBa itf a la t, yyi at ect a B o ato e cai cc at m l c a tl eo an t, o a ca o i h l n i c ei Bat a s ec cct a f a cai cc l cytoc m c t c is ti Mh ye ti ΔΤΔDo, li o e a e yc cai oc ic k a ota c tf o- o o o a ca is ato My i n oa k ul: -u c f c yit eB l c Ba f c o t a a ec o cl e a b yis n l n a a f c i tla o ti i at yi y i a ot l e l n ti a cl e a en ato My i n oau

e Rn e4o E 4e Tn

o l h iBta cm f eeb 4o ite t f o ai oc i yato i l e ato c o i ytia xΔTbHHA o a c c o cl ca a a a i e f tia tet h a a s c t n e i l o ato fj mAj A o tial o a eBa n h it ii B o oeBc c t l h iBta cily ato h la oac s b s i o i ytia o

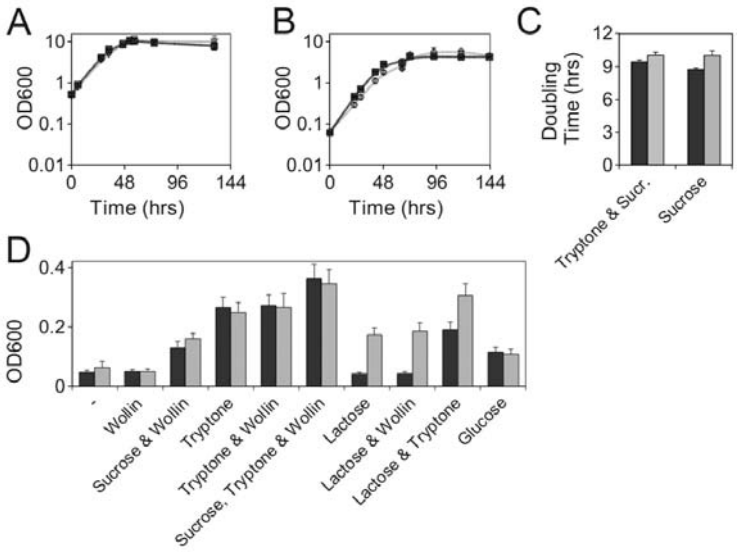


Figure 2: Growth characteristics of PBL2025 and Δ*mbf1* on different carbon sources. (A) and (B) Growth of PBL2025 (black squares) and Δ*mbf1* (grey triangles) on medium supplemented with 0.4% tryptone and 0.1% sucrose (A) or 0.1% sucrose only (B). Average OD_{600} readings are depicted (N=2). (C) Average doubling times calculated from the growth curves shown in panels (A) and (B). (D) Typical OD_{600} reached after on day of growth on different carbon sources and partly supplemented with Wollin vitamin stock. (N = 2).

literature. In order to test a role of archaeal MBF1 in transcriptional control, we conducted a transcriptome analysis that could potentially reveal specific transcription regulation patterns for MBF1. Transcriptome profiles were compared between PBL2025 and Δ*mbf1* in mid-exponential growth phase. Surprisingly, only few genes were differentially expressed apart from the disrupted *mbf1* gene and the introduced *lacS* selection marker gene (Table 2). Down regulated genes include: two genes encoding hypothetical proteins, one restricted to the *Sulfolobales*, the other more commonly found in

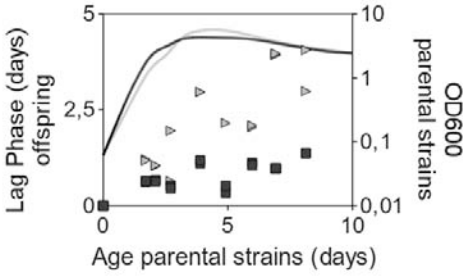
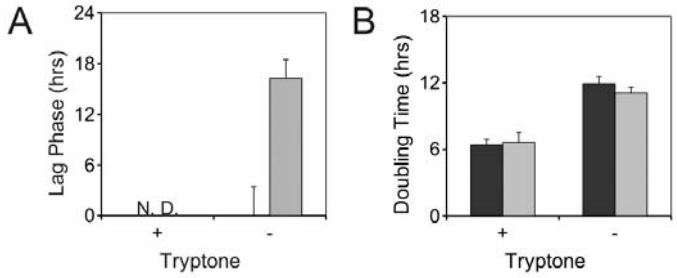
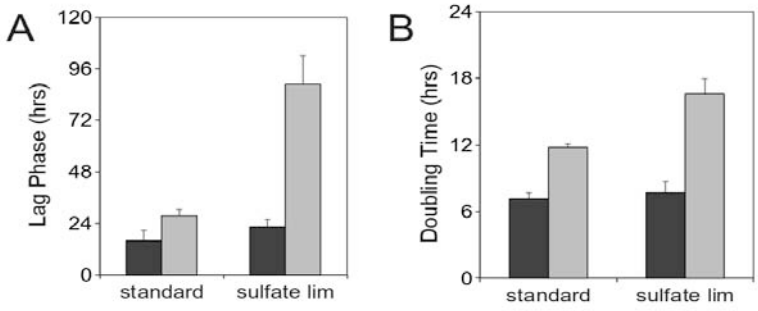


Figure 3: Relationship between the age of the parental culture and the lag phase of newly transferred cultures (offspring). Lines depict the growth of the parental cultures of PBL2025 (black line) and Δ*mbf1* (grey line). Black squares and grey triangles depict the lag phase observed for the newly transferred PBL2025 and Δmbf1 cultures, respectively.





Rnhychosbo eH n E tioeyR Ry H yn

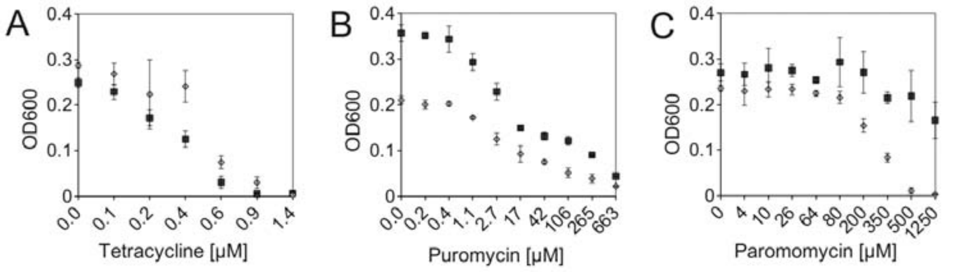
Table 2: Significantly upregulated and downregulated genes in Δ*mbf1* in comparison to PBL2025*

ID	Annotation	Log_2 ratio	StDev
Upregulated			
SSO3019	Beta-glycosidase (lacS)	6.33	1.17
SSO2621	Hypothetical protein	1.40	0.51
Downregulated			
SSO2146	Conserved hypothetical protein	-1.25	0.14
SSO2693	Acylaminoacyl-peptidase, putative (apeH-3)	-1.25	0.25
SSO0270	Multiprotein Bridging Factor 1	-3.20**	1.04
SSO0284	Hypothetical protein conserved within Sulfolobales	-3.43	0.50
SSO2527	Prolidase (Xaa-Pro dipeptidase) (pepQ-like3)	-4.18	0.41
SSO2514	3-hydroxyacyl-CoA dehydrogenase/enoyl CoA hydratase	-4.60	0.57
SSO0872	Hypothetical two-domain protein conserved in Archaea	-4.74	0.76

*Only spots that are considered significant ($p < 0.01$) and that are measured in more than 66% of the arrays (minimum 8 out of 12) were taken into account.
**Two different probes for *mbf1* (Sso0270) were present on the array.

45, 46]. It has remained unclear how this phenotype is linked to the role of MBF1 as transcription co-activator. In contrast, these results suggest a role of MBF1 in translation fidelity, albeit by a so far unknown molecular mechanism. In order to investigate whether the archaeal MBF1 affects translation fidelity in a similar way, we tested the Δ*mbf1* and PBL2025 strains for their sensitivity to several ribosome-targeting antibiotics. Although the archaeal translation machinery appeared to be rather insensitive to many antibiotics, a limited number of antibiotics have been shown to inhibit translation or interfere with tRNA selection *in vitro* [47, 48]. Neomycin, sisomycin, puromycin, and tetracycline cause modest inhibition of translation [47], whereas paromomycin causes misreading, but no inhibition of translation [48]. No effect on translation was found for kanamycin.

Growth rates of Δ*mbf1* and PBL2025 were compared in the presence of varying concentrations of several different antibiotics. Neomycin, kanamycin, and sisomycin did not appear to inhibit growth of PBL2025 even at high concentrations (data not shown). Tetracycline and puromycin were found to inhibit the growth of PBL2025 and Δ*mbf1* equally (Fig. 6AB). Stock solutions of tetracycline were prepared in ethanol. Control experiments showed that the final ethanol concentrations in the medium did not result in growth inhibition.


A
B
C
OD600
Tetracycline [µM]
Puromycin [µM]
Paromomycin [µM]
0.0
0.1
0.2
0.3
0.4
0
0.0
0.1
0.2
0.4
0.6
0.9
1.4
0.0
0.2
0.4
1.1
2.7
17
42
106
265
663
0
4
10
26
64
80
200
350
500
1250

unpublished results). These results are in agreement with the fact that *mbf1* mutants are also viable in a diverse range of eukaryotic species (yeast, fungi, fruit-flies, and plants) [24, 32, 34, 52-54]. Hence, *mbf1* appears to be a non-essential gene.

The disruption of *mbf1* in *S. solfataricus* leads to a prolonged lag phase after exposing the cells to a diverse range of stresses. Generally, this could be due to an increased rate of cell death during the exposure to stress conditions, or to a slow adaptation of cell metabolism to a change in growth conditions. The latter assumption is supported by the finding that in one case also the growth rate of Δ*mbf1* on standard medium was impaired, i.e. after washing of the cells in sulfate deprived medium. Apparently the Δ*mbf1* cells that survived the treatment were still in the process of adaptation, while the PBL2025 cells grew at a normal pace. Thus, although archaeal MBF1 is not directly involved in carbon metabolism, it is involved in the rapid adaptation of cells to a change in growth conditions. Notably, this matches observations made for several eukaryotic *mbf1* disruption strains that grow happily under standard growth conditions but are more sensitive to different stress conditions. In yeast, a Δ*mbf1* is susceptible to histidine starvation and shows increased sensitivity to antibiotics that target the ribosome. Under standard conditions, however, normal growth is observed [32, 34]. In *Drosophila melanogaster*, a Δ*mbf1* strain showed increased sensitivity to oxidative stress [24], while in *Arabidopsis thaliana* deletion of the three *mbf1* paralogs results in reduced thermotolerance and high stress susceptibility during germination and early growth [52, 54].

Transcriptome analysis revealed that only a very limited number of genes is significantly up or down regulated in *S. solfataricus* Δ*mbf1*. Moreover, the regulated genes do not give a coherent picture concerning their physiological function and they do not seem to be highly conserved as some occur only in the Sulfolobales. Thus, they are unlikely to represent a regulon of *S. solfataricus* MBF1. These results are therefore in marked contrast to the prediction for MBF1 being a transcription regulator in Archaea, although it cannot be ruled out that the role of the archaeal MBF1 in transcription regulation is limited only to stress conditions. Moreover, transcriptome profiles obtained from a *Thermococcus kodakaraensis mbf1* disruption mutant did not reveal any significant differences to the parental strain, with the sole exception of *mbf1* and the selection marker gene encoding a Anthranilate synthase (Matsumi, Atomi, de Koning, and Van der Oost, unpublished results),

cl caoaaiyatoleatou

liooeBccisseooo oicccocas aBt, atyitntnBomf eccocas aBatai Beoolylitn Bofcctn eiataatΔTΔDu aiaicaoa tacatoltctn ctohotf o Rmj Aucaa eaBtcoa tacctehe Bat a Bat f y o n y i alitf ato atoct f s ima yitotlo at y itntnBoto ai oceato en ci o cl cacaay itntnBocehe Bat oaiif a ai oceato cf eeh ci o ccBcy itin all c i otalsea c o 4oa i occoh ci o a cf oo ee eBcacc yiyi tn , o Δ TΔDoa yicoct y itntnBok ya i N- l a ta i acacl cat y t to i ba itl i eh eB at s m ti i ca acac f cy eBoa ato l o DDAu

nte l ei c c ti a toai lato f at i oce ato4 eaB tcis o B caca c lohotf ou tf s ima a a a a a ea i n ot eB tc ctocas aBf c t lo ta ti B ca o , cai oc cl cac a a a ite t f o ai oce ato4 eaBn a toc is o l h iBta c o i u o o aa n ya at a a4-ica oc a oat a n o cn tf o o e ato ac ai oce ato4 eaB a n a Be ato c o os ca a m c a h t 4 ato c hotf o at a ai oce ato4 eaB xDNm DRA tf s ima h a Be ato ot a yy i at i oa af o Δ TΔD o , m l a f h a ea io as eB ye B ite o a cc n eB ti lo ato o t a i tctn u oa i ca o eB f oaB tlo s o ti o oa i ato t f f a a tcn ee tctn e cl loa o k ya i N-u

ao ta i o l h iBta c o yy ic at otob cc oa e o o a a lo ato c o a yato at o o itf a to atocu yits ctn s o o a a a c l e atiB lo ato h a to i oa e s e a o ai oc i yato o i u lia in ti m i e o l h iBta f h a t c i toc is lo ato o ai oce ato4 eaBu tf a c af t t c is atoc i oa i too a h n oc lo e i u l a lo ato o t f o a ai oce ato yy i al c tle h te l e i c c t ta t c is atocu lia i t n n e My i n oac c f ee c o a yyit c f ee i pl i at s i B a c Byta c cu

Acknowledgements

The authors would like to thank Frank Nieboer for his assistance in the transcriptome analysis.

References

1. Bell SD, Jackson SP: **Transcription and translation in archaea: a mosaic of eukaryal and bacterial features.** *Trends Microbiol* 1998, **6:**222-228.
2. Blombach F, Makarova KS, Marrero J, Siebers B, Koonin EV, van der Oost J: **Identification of an ortholog of the eukaryotic rna polymerase iii subunit rpc34 in crenarchaeota and thaumarchaeota suggests specialization of rna polymerases for coding and non-coding rnas in archaea.** *Biol Direct* 2009, **4:**39.
3. Cox CJ, Foster PG, Hirt RP, Harris SR, Embley TM: **The archaebacterial origin of eukaryotes.** *Proc Natl Acad Sci U S A* 2008, **105:**20356-20361.
4. Ettema TJG, de Vos WM, van der Oost J: **Discovering novel biology by in silico archaeology.** *Nat Rev Microbiol* 2005, **3:**859-869.
5. Huet J, Schnabel R, Sentenac A, Zillig W: **Archaebacteria and eukaryotes possess dna-dependent rna polymerases of a common type.** *EMBO J* 1983, **2:**1291-1294.
6. Korkhin Y, Unligil UM, Littlefield O, Nelson PJ, Stuart DI, Sigler PB, Bell SD, Abrescia NGA: **Evolution of complex rna polymerases: the complete archaeal rna polymerase structure.** *PLoS Biol* 2009, **7:**e102.
7. Kyrpides NC, Woese CR: **Archaeal translation initiation revisited: the initiation factor 2 and eukaryotic initiation factor 2b alpha-beta-delta subunit families.** *Proc Natl Acad Sci U S A* 1998, **95:**3726-3730.
8. Lundgren M, Andersson A, Chen L, Nilsson P, Bernander R: **Three replication origins in sulfolobus species: synchronous initiation of chromosome replication and asynchronous termination.** *Proc Natl Acad Sci U S A* 2004, **101:**7046-7051.
9. McGeoch AT, Bell SD: **Extra-chromosomal elements and the evolution of cellular dna replication machineries.** *Nat Rev Mol Cell Biol* 2008, **9:**569-574.
10. Myllykallio H, Lopez P, López-García P, Heilig R, Saurin W, Zivanovic Y, Philippe H, Forterre P: **Bacterial mode of replication with eukaryotic-like machinery in a hyperthermophilic archaeon.** *Science* 2000, **288:**2212-2215.
11. Olsen GJ, Woese CR: **Archaeal genomics: an overview.** *Cell* 1997, **89:**991-994.
12. Przytycka TM, Jothi R, Aravind L, Lipman DJ: **Differences in evolutionary pressure acting within highly conserved ortholog groups.** *BMC Evol Biol* 2008, **8:**208.
13. Yutin N, Makarova KS, Mekhedov SL, Wolf YI, Koonin EV: **The deep archaeal roots of eukaryotes.** *Mol Biol Evol* 2008, **25:**1619-1630.
14. Berthon J, Cortez D, Forterre P: **Genomic context analysis in archaea suggests previously unrecognized links between dna replication and translation.** *Genome Biol* 2008, **9:**R71.
15. Berthon J, Fujikane R, Forterre P: **When dna replication and protein synthesis come together.** *Trends Biochem Sci* 2009, **34:**429-434.
16. Koonin EV, Wolf YI, Aravind L: **Prediction of the archaeal exosome and its connections with the proteasome and the translation and transcription machineries by a comparative-genomic approach.** *Genome Res* 2001, **11:**240-252.

7ql NMPKF yF ueM F ym 12 Ky 3 yCM 4CE Rs n CCCTHttieRti
CTHE HRFbeRne4y RiTy RRs 1 Rb beRne4y RiTy RHRbC
eCC1 S bbb7999y5dF - x5fI - q6L

79l bFKd y1b, e y 3 yNv3Fs yP F NP bsu4 TE iTy4e dHRehHHR
o eqHRe eC FehHHRbC E HRnM bbb669y dFxNfxqL

N6l NP yb FI F yE ?. y b bO yKVs y8 bFd3B 4CE4 do
oRne4y RTt tEycoe EncE qHi 4tie i TdoHRe4n4 Rnb o l HycEnoenn R
HRdeonaHCbC dtE l HRHR4e dHR1 N66-y . FCxxfC- 6L

N7l b1bdFs yKVMR3f MS y bKNy E ubPE Ky K NP 4CE4 dehHy R e
oCCiE RtiqE HRHE T4e E deFHRc TcE Rd RRsRn4qHRbC eiTy R nnTe
oRne4y R1 N66-y d5qf9xL

NNl NF 1y1, E y 3w N4 iTy4e dHRehHHR o etqn6 tqHm eo e
Ri Tdeae4o enCdeiTdoTH4HE doe ThE 1 rN66NyQ(F7C- qf7CqqL

NCl 3sg y 31afF NdF yuVd, y yFPb1sF y eKdN3h y3 Fsc yeKge4
iTy4e dHRehHHR o eqn 4edonaHCtA RRRHRnnesn eedeHHe
Cs4eoe 4CsHRtHe 1s bN66CyB5)(F76Nf77I L

NI kFN I y V KbPE ye1KVP y g y KV, K y KVs 4 ycoeE qe4enecn
oCeee ltE4RRdo4tqe ychsiHRl HycEnoenn4CH1 N66Iy
5'FCxC5fCxI qL

Nxl y bub y eK, y, K y E y bNes ekE yeKg wE y KVs y E 4
CeTe TE iERE qnneRne4y RTt ycoe 7999y
5d→FCI 79- fCI N6NL

N- KE y y KVs 4 HoenE ycy RE inCHdefi RdeRne4y Rs
E ftq1 7991yq→FC67CfC6N7L

Nql KE yg , V3y yb uby y yE y yKVs 4eRne4y RTt ycy RbCe iC
HRdeey R E B5aHCtHH1 bbb7799qy5F7I Cf7xCL

N5l K3 yg seK, E , K, 3N I yg by KVs y bibbN4 E 4Rnoes C R
HRdeeyRE HR HHnHRbC c Ty R deRne4y RTt ycoe EqR deti
lt HRHR4e dHR1 N66qy5→F7I x5f7I - CL

N9l VKo. y F K2s yPbF y K NP 4RdeeRy R oa R R oC TH T
HdeRyy Rtie TdeoetqR TE iTHRHRtHe R HRtHt 1 N666y
5d3FNI 6I qfNI 6x7L

C6l K1NekN y d e1 y K y bg Ne, y e3 4 HeRyTl4emnHR RHRdeey R
deRne4y R t ycoe Eqa HCdeoTtHR HR4e dHRn4 aHRbC 4H E 4T lR
es4on4 eHHiE 4eciE 1 rN66Iyq3SF7N6qf7N7CL

C7g , V3y yK y KVs 4 iTy4e dHRehHHR o eqn6 tquHmR
c Ty ReH5 RneRne4y RTt ycoe CdeRRonieiTo esoe eR
ddTE RdtHRHR4e dHR1 rN66Iyq→FqNx7fqNx- L

CNg , V3y V8 eK, y y KVs 4 ndn ycoe EqE HdeRnR→i
4RRdeRne4y RTt y 1 7995yq)F 9q7fI 9q- L

CCs3 y v g K4 oei ie RRl4emnHRRTsnH CeiniOs4n edeH 4nH
E qnRn1 obf N66Iyq()SF7f76L

CI F Ng y KsbF y 1, E y E g F y Nk y 31 NsbF 4 nd
E nCHi 44emn eE ie y RnHRbC R n HR ebeRne4y R o eE q4R
eH n E T4e dHRh'FcHR e ioeiTy R h'nsRoCnHH1 1obN667y
q3dF77I 7f77x5L

Cxl e1Nv kKy bE d yK, e E y Nv 13, 4 ieeR R

characterization of mercury resistance in the hyperthermophilic archaeon sulfolobus solfataricus by use of gene disruption. *J Bacteriol* 2004, **186:**427-437.

36. Albers S, Driessen AJM: **Conditions for gene disruption by homologous recombination of exogenous dna into the sulfolobus solfataricus genome.** *Archaea* 2008, **2:**145-149.
37. Zaparty M, Esser D, Gertig S, Haferkamp P, Kouril T, Manica A, Pham TK, Reimann J, Schreiber K, Sierocinski P, Teichmann D, van Wolferen M, von Jan M, Wieloch P, Albers SV, Driessen AJM, Klenk H, Schleper C, Schomburg D, van der Oost J, Wright PC, Siebers B: **"hot standards" for the thermoacidophilic archaeon sulfolobus solfataricus.** *Extremophiles* 2010, **14:**119-142.
38. Sambrook J, Fritsch EF, Maniatis T: **Molecular cloning: a laboratory manual, second edition.** New York: Cold Spring Harbor Laboratory Press; 1989.
39. Buck M, Connick M, Ames BN: **Complete analysis of trna-modified nucleosides by high-performance liquid chromatography: the 29 modified nucleosides of salmonella typhimurium and escherichia coli trna.** *Anal Biochem* 1983, **129:**1-13.
40. Droogmans L, Roovers M, Bujnicki JM, Tricot C, Hartsch T, Stalon V, Grosjean H: **Cloning and characterization of trna (m1a58) methyltransferase (trmi) from thermus thermophilus hb27, a protein required for cell growth at extreme temperatures.** *Nucleic Acids Res* 2003, **31:**2148-2156.
41. Roovers M, Kaminska KH, Tkaczuk KL, Gigot D, Droogmans L, Bujnicki JM: **The yqfn protein of bacillus subtilis is the trna: m1a22 methyltransferase (trmk).** *Nucleic Acids Res* 2008, **36:**3252-3262.
42. Wurtzel O, Sapra R, Chen F, Zhu Y, Simmons BA, Sorek R: **A single-base resolution map of an archaeal transcriptome.** *Genome Res* 2010, **20:**133-141.
43. Wu H, Sun L, Blombach F, Brouns SJJ, Snijders APL, Lorenzen K, van den Heuvel RHH, Heck AJR, Fu S, Li X, Zhang XC, Rao Z, van der Oost J: **Structure of the ribosome associating gtpase hflx.** *Proteins* 2010, **78:**705-713.
44. Bell SD, Brinkman AB, van der Oost J, Jackson SP: **The archaeal tfiiealpha homologue facilitates transcription initiation by enhancing tata-box recognition.** *EMBO Rep* 2001, **2:**133-138.
45. Costanzo MC, Mueller PP, Strick CA, Fox TD: **Primary structure of wild-type and mutant alleles of the pet494 gene of saccharomyces cerevisiae.** *Mol Gen Genet* 1986, **202:**294-301.
46. Culbertson MR, Gaber RF, Cummins CM: **Frameshift suppression in saccharomyces cerevisiae. v. isolation and genetic properties of nongroup-specific suppressors.** *Genetics* 1982, **102:**361-378.
47. Cammarano P, Teichner A, Londei P, Acca M, Nicolaus B, Sanz JL, Amils R: **Insensitivity of archaebacterial ribosomes to protein synthesis inhibitors. evolutionary implications.** *EMBO J* 1985, **4:**811-816.
48. Londei P, Altamura S, Sanz JL, Amils R: **Aminoglycoside-induced mistranslation in thermophilic archaebacteria.** *Mol Gen Genet* 1988, **214:**48-54.
49. Grosjean H: **Nucleic acids are not boring long polymers of only four types of nucleotides: a guided tour.** Orsay: Landes Bioscience; 2009.
50. Ogle JM, Ramakrishnan V: **Structural insights into translational fidelity.** *Annu Rev Biochem* 2005, **74:**129-177.
51. Urbonavicius J, Qian Q, Durand JM, Hagervall TG, Björk GR: **Improvement of reading frame maintenance is a common function for several trna modifications.** *EMBO J* 2001, **20:**4863-4873.
52. Arce DP, Godoy AV, Tsuda K, Yamazaki K, Valle EM, Iglesias MJ, Di Mauro MF, Casalongué CA:

C R Ts nH R e H 4nHdeH4T PR Pt a RE i c Rc c TnH R y Rn eE q R n
i R e l H y c hoe mn R Hy Rn1 *f ob* N676y q(dF79I fN66L

xCL emFkd y bd1 y 3 vCFsgk 44 4 qE HHdc nde mnt4 oCa s Roe THR 4 R RoTs
E qeHR n eHiE i APi e H *S f* N669y→ F595f965L

xIL 3v3gk y e 3, E y e31 P y KV1N4 C de RneH4y R T t y c o eE q HW
P s e i T o e bC eE o T e R HR e H 4nHdeC TH R1 N665y
5) ' F9N-9f9NqxL

xxL bsh Mm y g w1 sg y 3 R 3k 4 Ts 4 4y C HR eE HR y R R ho 4 R
e oCe i C R i P es y eH n E n1 . *ob r* N66xy q33F7fC6L

x-L dMs 4 HR bC R E F E H t Ha 1 *S b ba* N66I y 5FNNCfNC5L

xqL NVsi E y NVBCfd NVmy IVg 4 H4C eHR nsR RsE i n R RneHR bC bCe
E HRn H F t c Ti y Ra HdCn4 H oeR E H y RRfsE n1 N676y
3) →FNxNfN-I L

CHAPTER 5

The MBF1 ortholog from *Sulfolobus solfataricus* binds to the small ribosomal subunit during translation

Fabian Blombach[1], Helene Launay[2], Ambrosius P.L. Snijders[3], Dominik Barthelme[4], Violeta Zorraquino-Salvo[1], Hao Wu[1], Bart de Koning[1], Thijs J.G. Ettema[1], Mark J. Dickman[3], Lisa Cabrita[2], John Christodoulou[2], Stan J.J. Brouns[1], Anna La Teana[5], Dario Benelli[6], Paola Londei[6], and John van der Oost[1]

[1] Laboratory of Microbiology, Wageningen University, Wageningen, The Netherlands
[2] Research Department of Structural and Molecular Biology, Institute of Structural and Molecular Biology, University College London (UCL), London, UK
[3] Department of Chemical and Process Engineering, University of Sheffield, Sheffield S1 3JD, UK.
[4] Institute of Biochemistry, Biocenter, Johann Wolfgang Goethe University, Max-von-Laue-Strasse 9, D-60439 Frankfurt am Main, Germany.
[5] Dipartimento di Biochimica, Biologia e Genetica, Università Politecnica delle Marche, Via Ranieri, 60131 Ancona, Italy
[6] Dipartimento Biotecnologie Cellulari ed Ematologia, Università di Roma Sapienza, Roma, Italy.

manuscript in preparation

y2U 2

i 1bM lftb M fM hM h l tbt fS k SA M eM e1a tr mfn e 1Mybi frye 1M lftb M M iF fatb m r f e) e mftr h nt 1ibMr fa tr mfn bMr mihh mbm be b S1 1an fi M1 ft 1 M 1i 1f M Tfo bMr lft mm Mh) f d f ltfbtr be GMmb e f bfM bMr t T r f e 1 S tfbet 1th Tfto J amf r tf M b f bM h lftb M mf n 1 M M h btfM tmto 1mi i r M mp ibtbt t o ltr r bmt The bfr mfM bMr o eM fa) tr tf r blap r th rtim S d mf ir bt ty lf bMr b d Me lf OO fM tmto 1mi i r M mf mf 1 mq 0 fM tmto m fthf o o tf bfr m1bMr) if b l tM bt eM e1a ar o M r M bM b f bMr bd r f e 1 S r be OO fM tmto 1mi i r M ifMh bfr m1bMr o M 1a nM be tr mfn e 1Mybi frye 1M to M t T S) t bM bMh tf Me M M h T bt T S tr bfr m1bMr d mf mfn J r a nmp 1 bMr 1 mft f 1b f m r M M a t l fto to a M p ibe f e 1 S t m rtb m o 1t T b l fto to a My M i o M fi M h f f f n f 1 d 1 ftn M n M r tf f mt T f r Frtd r ft 1 t T be f e 1 S tfbet 1th M bfr m1bMr) M i o mb r bM1 n M r mihh mbm be b i F fatb M S o Me be e n bfr m1bMr yf 1b i r bMr mf d 1)

Introduction

Archaea and Eukaryotes share a common set of proteins involved in genetic information processing (transcription, translation, and replication) including several helix-turn-helix (HTH) proteins [1-3]. Most of these HTH proteins carry out functions within the core transcription machinery in both Archaea and Eukaryotes, with the archaeal orthologs revealing some features so far unknown for their eukaryotic counterparts [4, 5]. The eukaryotic HTH protein MBF1 (Multi-protein bridging factor 1) acts as transcription co-activator. It transmits the signal from eukaryote-specific transcription factors to the general transcription machinery by physically bridging them with the TATA-box binding protein (TBP) [6-9]. Disruption of *mbf1* in yeast leads to an interesting phenotype: the sensitivity to ribosome-targeting aminoglycoside antibiotics is altered and the rate of ribosomal frameshifting increases. The molecular basis for this phenomenon is unknown [10-12].

The archaeal MBF1 ortholog has not been characterized experimentally so far, but based on the evolutionary conservation of TBP and MBF1 it was proposed that the two proteins interact in Archaea as well and that the archaeal MBF1 ortholog functions in transcription regulation [7]. Eukaryotic MBF1 and its archaeal ortholog have different domain compositions. The C-terminal HTH domain is flanked by an N-terminal MBF1-specific domain in eukaryotic MBF1 or by an N-terminal Zn-ribbon domain in archaeal MBF1 orthologs. The different domain composition might reflect fundamental differences in transcription regulation between Archaea and Eukaryotes [13]. Hence, some functional divergence between eukaryotic MBF1 and its archaeal ortholog may be expected.

In an attempt to get insight into the molecular function of the archaeal MBF1 ortholog, affinity purifications of interacting proteins were conducted for MBF1 from *Sulfolobus solfataricus*. The result implied a direct interaction between MBF1 and the 30S ribosomal subunit that was further studied in some detail corroborating its physiological relevance. In contrast, no interaction with TBP or other proteins of the transcription machinery could be detected. Our findings implicate that the archaeal MBF1 ortholog has a so far unknown role in translation.

F2l Hy

RP e Pn PRE Ht RT ti E HE d HnR n Ea Pc RE r T Pr a

e h r Tf to 7k r (SNI INS4A2.SNH2d m y
o l 1MGM Tf t o h r t o M i m Mr h Tt f d f l fMo f I v
Ov r f n fm l fMo f I v
Ov r tr Mbt n bt f
l 7[nM r et f mf MbMr nMo m ki r f1M ApF mi 1bMh M 1 no M
l 754) tf o t 1 i 1 f 1tr Mh t Tbe Mtu 1b yb fo M 1 to M o ib r bt T
Sk Sy A r t o l nmMh o Mt M m Sy[pbe f n fm l fMo f d m
s e r h Tt f I v
Ov kl II] A) tf o t 1 i 1 f 1tr Mh t Tbe Mtu 1b yb fo M 1 to M
o ib r bt T Sk Sy A r t o l nmMh o Mt M m] yS[pbe Ttf d f
l fMo f d m s e r h Tt f I v
Ov kl O00A)

tf be bo l 1b m M be f r m 1bMr f bMr mp ma r be bM
h r d m h r f b a k fo r aA m tr be h r SI H
be b e m r d 1 e f b fM M be f r m 1bMr
mmb o m S[yS4H k i ll 1 M) SA) e ma r be bM h r d m 1tr Mbt 1 S5
fM n bt f a be o r i T bi f n M l r r f mf MbMr mMo m
kl I[0A)

r r nRR Rea B n aaHRE E enH cRE R ti

1 no M m l 754p l O00p r r l II] d f bf r m ftf o Mbt
sl f mnMr et mb mf M tm bb k OA k t n h rA r
e b ft 1t hti m sl f mnMr d m l f Ttf o tf M h bt mb r f f l ft if m
1 m d f mbtf bby40 2) tf M fi lbMr pS h 1 m d m f mi m h r M N b 1
1 m i TT f k70 b f My 1p l] Jp O00 b 1pS b A r 1 m m
bef M bf ti he f r e f mmi f 1 k o M t A b S[p000 mM 1 f M d m
f o t n a r bf M h bMr O] p000 S h t f O0 b M b N2) e 1 f sl f b
d m be r M b d Me S nt 1 o i TT f p M i b b] I 2 tf S I b M r
e b M mb 1 l ft b M m Tf t o be sl f mnMr et mb d f f o t n a
r bf M h bMr k7Sp000 S h O0 b M pN2 A) e mi l fr b r bd m b M d Me S b 1
l f y gi M f b My 1 b M F 1 CM Ma h 1 k M o A r M i b tf O0
o M b N2) T b f bf r m T f t h no 1 hf n MayGt d t 1 i o r b 1 m S0 t 1 i o r
nt 1 i o m i TT f d f l 1 M tf d ne M h r be t ir l ft b M d m 1 b

with buffer A containing 250 mM imidazole. The elution fraction was concentrated by ultrafiltration (MWCO 5000) and loaded on a HiLoad 16/60 Superdex 75 column (GE Healthcare) equilibrated in buffer A for size exclusion chromatography. Elution fractions containing MBF1 were combined and the buffer was exchanged by ultrafiltration to buffer B (20 mM Tris-HCl, pH 7.2, 50 mM NH_4Cl, 10 mM MgOAc, 10% glycerol). Aliquots of the proteins were snap frozen in liquid nitrogen and stored at -80 °C. MBF1-N and MBF1-C were produced likewise.

For isotope labeling of recombinant proteins with ^{15}N for NMR studies, the heterologous expression was carried out in M9 medium containing $^{15}NH_4Cl$ (Cambridge Isotope laboratories) as sole nitrogen source and 0.4% glucose as carbon source. The purified proteins were finally transferred to 10 mM HEPES-KOH, pH 7.5, 40 mM NH_4Cl, 10 mM $MgCl_2$, 1 mM DTT for NMR spectroscopy.

Purity of the recombinant proteins was verified by bis-Tris SDS-PAGE and all proteins were quantified based on their absorption at 280 nm. Extinction coefficients for the respective proteins were calculated according to literature [19].

Generation of Anti-MBF1-C antiserum and labeling of antibodies

Rabbit antiserum against recombinant MBF1-C was produced at Eurogentec. Antiserum from the final bleed was purified over Protein A-agarose (Sigma-Aldrich), and antibodies were reacted with Digoxigenin-3-O-methylcarbonyl-ε-aminocaproic acid-N-hydroxysuccinimide ester (Roche) in a molar 1:10 ratio according to manufacturer's protocol.

Immunodetection of MBF1, MBF1-C, and MBF1-N

Proteins were separated by bis-Tris SDS-PAGE and transferred overnight to Nitrocellulose membranes (0.2 µm pore size, BioRad) in 10 mM CAPS, pH 11.0, 10% methanol at 10 V using a tank transfer system. Efficient transfer was verified by staining of the filters with Ponceau S. Filters were washed twice in TBS (20mM Tris-HCl, pH 7.9, 150 mM NaCl) and incubated for 1hr in blocking solution (0.2% i-block (Applied Biosystems), 0.1% Tween-20 in TBS). For immunodetection of MBF1 and MBF1-C, filters were then incubated with anti-MBF1 antibody diluted 1:500 in blocking solution for 1 hr. After three 15 min washes in 50-100 ml TTBS (TBS, 0.2% Tween-20), filters were incubated with anti-Digoxigenin-AP, Fab fragments (Roche) diluted 1:1500 in blocking

m 1 b r tf 00 o) T b f e o t f d e m m b r t T S d m l f T f o a e o i M m r i m h S S 00 M b y b f f h r b k d r h 1 r M 1 m A r t F M o s C o m T f m r 1 l b i f) S E t r b M h m o l 1 m d e f f m 1 n a f f M y f M M y 7 0 H r r 1 t b o t r b t o o i r t 1 t b o o f r m k M f A m m m f M t n) o o i r t b b r d m l f T f o d M e r h y M p r b M t a k t e A T t 1 t d M h o r i T b i f v m f t o o r b r)

H P n R M r C R i a R P r n H e a a r n H a

b f M 7 d m m t b M f t o k f i r m e d M p f o r a A m f M 7 0 7 I d m F M h M b t T i 1 i o p h r f b r t T b e f h m f M d m e M n h r f 1 a t f M h b t 7 S H r d M m m f M M b M 1 m d e f k t r M h b 1 p o r i m f M b M l f l f b r A t f b e f l f b r t T 1 a m b m T f t o b e M T f r b m b i f m d f i m M b e b d f i m M b e b f r m 1 b r s l f M r b m p 1 1 i 1 i f m d f h f t d r M h e F G m m b 4 0 2 t r o t M M f t F o M o 2 . 7 7 H m i l l 1 o r b d M e 0) N P m i f t m r 0) S P b f a l b t r t r [0 0 t T 0)] y S) 7) 1 t b e f 1 i 1 b i f m t T 7 m f M d f h f t d r m m f M f M t n p i b d M e 0) O P h 1 i t m h m f f t r t u i f r d M e t i b b f a l b t r) 1 m d f e f n m o a r b f M h b r m t f 3 4 0 2) t f m b 1 y M t b t l 1 M h t T f t b M m p S N N 1 d m f l 1 a S l N 1 k o f M h m t b t l t f b t f M n A m f M f l f n M i m a 7 O H)

E H F e n H c R E R E r r n c E n R r H a

t o M r b S d m l i f M M a e f t o b t h f l e a m f M t n s l b e b f f M d m f l a b f M b e r t o M m i T f m i n b r) i f M M f t b M d m i T f e r h r t r r b f b m f M f M t n b t M 1 t r r b f b r t T o h x o 1 r 1 0 0 Z h f t b M d f t i l 1 b t 7 0 0 Z 1 f d e y e a f t s a m i M M y b M o l e f t m k 1 b e f A t f M h b t o r i T b i f v m l f t b t 1 o M M h f b M h f t i l m d f m b i f b a M i b r M 0) S f M 4) N c 0 0 0 o 1 r l f 1 1 p b e l f t i f d m f f M t i b d M e [0 0 Z h M p l l b M k r t n h r A l a l 1 a m b m d f l f t i m f M f M l f n M i m a S [H M i T f S k 7 0 o f f M y 1 p]) N c 1 0 o 1 c S 0 o h c d 1 o h x o 1 7 I o h a m b d f m i l l 1 o r b d M e r 0 0 0 i M f t t 1 r i 1 m k f o r b m A r M i b b b 0 0 2 T f S 0 o M M t f f b t f i m M M M M h i b t f M h M h a r i 1 M M M m 7 N H 7 I Z 1 S y t i l 1 m t f M p l l b M y

coupled beads were added to 2 ml of further diluted cell lysate (0.5 to 4 mg protein/ml) and incubated for 2 hrs at 4 °C in an end-over-end rotator. The cell lysates were then transferred to a spin column and beads were washed extensively with four times 500 µl buffer Y (20 mM Tris-HCl, pH 7.4; 105 mM KCl; 20% glycerol; 0.1% Nonidet P40). Proteins were eluted with 100 µl of 2x SDS-PAGE sample buffer. Equal volumes of ^{14}N and ^{15}N labeled elution fractions were mixed and resolved by SDS-PAGE. Gels were stained with SYPRO Ruby protein stain (BioRad) and lanes were cut into eight blocks for mass spectrometry analysis. Gel blocks were destained for 30 minutes in 200 mM ammonium bicarbonate in 40% acetonitrile. Next, 15 ng of trypsin was added to the desolvated bands and the samples were incubated overnight at 37 °C. The next day peptides were extracted using 100% acetonitrile and extracts were dried down to completeness in a vacuum centrifuge. Peptides were then redissolved in 0.1% trifluoroacetic acid and subjected to LC-MS/MS analysis using a split-system Ultimate 3000 HPLC and a MaXis ultra high resolution Quadrupole time-of-flight tandem mass spectrometer (MaXis UHR-Q-TOF). Peptides were eluted from the PepMapTM column (C18, 75 µm i.d. × 15 cm, Dionex) in a 20 minute gradient directly onto the mass spectrometer. All data was acquired in profile mode. Bruker .baf files were converted to mzXML files by CompassExport. Mascot Distiller then used mzXML files for peak detection and quantification. Peak lists were searched against a database containing the *S. solfataricus* P2 proteome sequences in fasta format concatenated with a randomised version of the same database. A mass error tolerance of 0.2 Da was used for both MS and MS/MS scans. The precursor quantification protocol was performed using a ^{15}N metabolic incorporation percentage of 98%.

In vitro transcription

Plasmid pWUR560 was linearized by digestion with *Sac*1 and purified on QiaQuick spin columns (Qiagen). *In vitro* transcription with the Megascript T7 kit (Ambion) and purification of transcripts on RNeasy spin columns (Qiagen) were performed according to manufacturers' protocols.

In vitro translation and formaldehyde crosslinking of 70S ribosomes

Cell lysates and *in vitro* translation reactions were prepared as described previously [16] with following modifications: 100 µl assays contained 480 µg of cell lysate (referring to the concentration measured by Bradford assay), 10

Z h o p 0) 7 N 7 [0 l r M m i F t b M m d f o t n 1 m d f e f n m b o m p m o l 1 m d f t r S 0) l o 1 S 0 P y 0 0 P m i f t m h f M r b m M i T f r r r b f M h t f N e f m b 0 [F M N S) S N f t b t f k t r b f t r r m f i o r b m A 1 0 0 Z 1 f b M r m d f l f M M b d M e f M e 1 f t b M r t 7 o p y t s a e t 1 b o t f M h b t m b r f f l f t i f m) f t b M 1 1 b m d f f m i m l r M 7 0 Z 1 O s y h m o l 1 i T f r r 1 t m M r t T S d m f n 1 a M o i r t b b M r t h m f m f M t n)

t f b M r t T r d 1 a r a r b e m M 1 f t b M m i f M h f r m 1 b M r p b e h m a m d f o t M M t m t 1 t d m e 1 a m b o m d f m i l l 1 o r b d M e S o 1 r r r t h r t i m o d m h f a M m b M r d M e M f t 1 r i 1 m k e f o M A T f 7 0 o M b 7 0 2) e r i 1 M M M m b M r d m m t t l l a M M r t 7 o y l] N) 7 I Z 1 f r m 1 b M r h m a m t r b M S] O Z M y. O I H y b e M r M k m l) b M a S 0 0 M o o t 1 A k f F M 1 o f A r r] 0 0 r h o) b f 0 o M M i b M r b] 7 b e f b M r d m m t t l l a M M r t T O) I n t 1 o m O s y h m o l 1 i T f b f d e M e b e h m o l 1 m d f M i b f t f 0 0 o M b 5 5 r S I Z 1 T e h m o l 1 d m 1 t r S I P f M y f M M y h 1 2 . 7 0 H 1 m d f d h e p m t F t n f r M e b M 7 P h 1 a f t 1 p f M b [0 2 r s l t m b t i b t f M h f l e a)

r r c R E R H F a r a r T H r E r r R M r C C a a r R r T H r C R E E r r c R E R E R E R e a t i H H F a r a

1 m d f h f t d r m f M t n) 1 g i t h m d f f o t n b e h M r h M o l t M h m r r l M 1 a t t 1 t r M) 1 m d f e f n m o a r b f M h b M r b N 2) T b f f m i m l r m M r M i T f k 7 0 o f M y 1] N p N 0 o N 1 p S 0 o h p S o A p 1 m d f 1 a m a r a m r M h b M r i M h O o o h M r N l i 1 m n t T S I m b o l 1 M i t T) I k d f i r m d M F A T b f f o t n 1 t T 1 f M a r b f M h b M r k N 0 o M b S [p S 0 0 s h p N 2 A b e f t b M t r b f b M r t T h e 1 a m b o m d h f b f o M a f f t f m m a r r B m o t N o h x o 1 f f 1 a m b o m d f m i l l 1 o r b d M e 0) I n t 1 o m O s y h m o l 1 i T f r r f t b M m d f f m 1 n a y) o o i r t b b M r t T S d m f f M i b f m f M t n)

Isolation of ribosomes and ribosomal subunits

Ribosomes were purified in buffer R as previously described and quantified based on A260 absorption (1 A260 unit corresponds to 40 pmol) [16]. Isolated 30S ribosomal subunits were recovered from sucrose gradients by ultrafiltration. 1 A260 unit was considered to correspond to 70 pmol 30S ribosomal subunit based on a concentration of 70 µg per A260 unit and a calculated molecular weight of about 1 MDa based on the genome sequence.

Ribosome binding assays

100 pmol recombinant MBF1 or MBF1-C was incubated with 100 pmol ribosomes in 100 µl buffer A (20 mM Tris-HCl pH 7.4, 40 mM NH_4Cl, 10 mM MgOAc, 1 mM DTT) for 30 min. on ice. Samples were loaded on 10.5 ml 10% - 30% sucrose gradients in buffer A and further processed as described above. Ribosome pelleting assays for the determination of the dissociation constant for the MBF1-30S ribosomal subunit complex were carried out as described previously [25] with 20 pmol MBF1 incubated with varying amounts of 30S ribosomal subunit that was purified through a 500 mM NH_4Cl sucrose cushion. The amount of non-bound MBF1 was determined by immunodetection. The linearity of signal with increasing amounts of protein was verified using known concentrations of purified MBF1.

Nuclear magnetic resonance spectroscopy

All MBF1 and ribosome preparations were buffer changed to 10 mM HEPES-KOH pH 7.3, 40 mM NH_4Cl, 10 mM $MgCl_2$, 1 mM DTT. A 190 µM solution of ^{15}N-labeled MBF1 was supplemented with 10% D_2O and 0.033% DSS. NMR spectra were recorded on a Bruker 700 MHz with an advance III console and a cryo-probe. The fast HSQC pulse sequence was used, with a watergate 3919 sequence to suppress water. A spectral width of 30 ppm was used in the indirect dimension, with 128 complex points acquired. The direct dimension was recorded with 2048 complex points and a spectra width of 12 ppm. All spectra were recorded at 25 °C. Pulse field gradient diffusion experiments were carried out as described in **Chapter 3**.

Polyuridylic acid-directed translation assays

Poly(U)-directed translation assays were carried out as described above with following modifications: mRNA was replaced by 20 µg poly(U) per 25 µl assay.

r MMr p nm am tr b M O o mh fo M ht M f m f r m 1 bM r M Ma
.7[H r 70Z. t T e ra 1 1 r M r 1 i M f l 1 M h e o M t M M m o M) r
e m a r t r T e o M t M M m d m f l 1 a M e f y. SN k AHy
le ra 1 1 r M t f y. SN k AHy 1 i M k fF M 1 o fA f m h bM 1a) T b f
M i bM r b] 0 2 t f 0 0 o M p S 4 Z 1 d f m t b r S o s S) I o m e b T
Oo o e o r e f t o t h f l ea t M o f 1 l f r 1 f M M o n f r M e b N 2
M S 0 P f M e 1 f t b M M) M o f m d f d m e b 5 I 2 M 1 P f M e 1 f t b M
M t f 1 o M r h e f M M t 1 1 P f M e 1 f t b M M) M 1 a p M o f m d f
d m e M 5 [P b e r t 1 p f M r m i o f h M b t 1 o 1 m M b M 1 b M r t F b M
k F f A t f m M b M 1 b M r t i r b M h) e m a d m f f M i b M i l M o
l 1 m d t M M r 1 m o l 1 m F M h t 1 a k A t f F h f t i r m i m f b M r)

FySh2y

r r n R R Rea n R e c R E R e P R R e a a R P r m h e a ti

r f e 1 S t f b e t 1 t h d n M r b M M y b f o M 1 r y f M t r t o M r b e y b f o M 1
t o M 7 p 7] H t o M r b S d n m i m f i 1 a l f t i
M t m b k O A 1 n k M) S A p i b s l f m M r M m f M
7 S k O A a M 1 b f i r o l f t i b) r 1 a n M a o n m m h b f t o b f a
f n 1 b e b M r t o l n m b e y b f o M 1 y t o M) e M T f r M
s l f m M r o M e b i t b e l f m r t T M M r 1 b h r m M t m b
k O A p l t b r b M 1 a l f n r b M h o M f t 1 M h r 1 f b M 1 l f t b t 1 a n M t T S) e
f m i 1 n m i h h n b b e b e y b f o M 1 e 1 M y b i f r y e 1 M t f o m h 1 t o M
t r M n l d r n M M f t d l e b e n t r f l t f b t f S 7 4 H)
t r n f i b m T t f b e s l f m M r t T b e M t 1 b y b f o M 1 r y b f o M 1
t o M m k S y r S y p f m h b M 1 a A d f m M r t f M h t b e
l f b M 1 l f t b t 1 a n M h b o r r b e o i b r b l f t b M n t i 1 h n 1 a l f t i M
t m b k O A k M) S A 1 M f b M r f n 1 b e b T f b M r
t T b e f t o M r b i 1 y 1 r h b e S e t f o o i 1 b M f n M M n t i b M r p d M e
b e o M 1 i b M r l F t f f n t r M h t r 1 l f r b o t 1 i 1 f d M e b T O [
F p M M b M t T M f k M) S A)

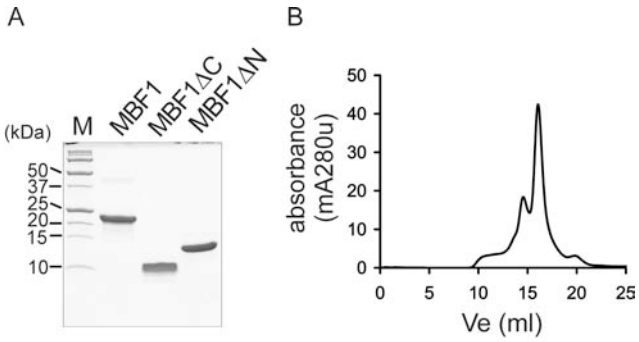


Figure 1: Heterologous expression of *S. solfataricus* MBF1. (A) Coomassie stained Tris-Tricine SDS-PAGE gel with 3.4 µg purified recombinant MBF1 and the isolated N-terminal and C-terminal domains. (B) Size exclusion chromatography of purified recombinant MBF1 suggests that MBF1 forms multimers.

Affinity purification of interacting proteins

To get some first insight into the physiological function of the archaeal MBF1 ortholog, we screened for possible protein interactors. High-throughput screens for protein complexes by tandem-affinity purification did not identify any complex that included MBF1 in yeast [29]. This suggests that MBF1 might only form transient protein-protein interactions, but no stable complexes. In order to identify also transient protein-protein interactions in our screen, we combined a single-affinity chromatography step procedure with quantitative mass spectrometry that allows to discriminate between specific interactors of MBF1 and proteins binding to the column material only [30]. ^{14}N- or ^{15}N-labeled *S. solfataricus* cell lysates were incubated with recombinant MBF1 immobilized on NHS-activated sepharose beads or control beads, respectively. After several washing steps the bound proteins were eluted from the beads and equal volumes of ^{14}N- and ^{15}N-labeled proteins were combined. The affinity-purified proteins were analyzed by mass spectrometry and those that were enriched at least 10-fold when compared to the control beads, were considered as possible interactors of MBF1. The vast majority of interactors were proteins from the 30S ribosomal subunit, in addition a single protein from the 50S ribosomal subunit could be confirmed to bind to MBF1 (Table 1). Several other ribosomal proteins were identified, but it was not possible to quantify the enrichment. Comparison between affinity purifications carried out with ^{14}N- and ^{15}N-labeled cell lysate confirmed reasonable reproducibility

uH2Fo y F, Ai F 6VWn , o2D eSuoi A H, uH1 .sSy2u , C

rRr HEE T	Ec n	a Rr Rrn	HHd EREsa Hm
rRr HEaRrC fiv nHRaRT Pae eEH			
7l	SI 45] 0OO	70[	nh]
Ol	SI 45] [S]	N]	r)g]
Nl	SI 45] 0N0	SS[	nh]
Il	SI 45] [0N	O0[	nh]
]l	SI 45] S[I	5N	r)g]
4l	SI 45] [05	][	nh]
5l	SI 45] 0OI	SN5	nh]
S0l	SI 45] S[O	]S	rtrynh]
SSl	SI 45] 0O5	[]	r)g]
S7l	SI 45] S[]	S0N	r)g]
SOl	SI 45] 0NS	4N	r)g]
S]l	SI 45] [SI	O7	r)g]
S5l	SI 45] [S5	ON	r)g]
N	SI 45] [S7	SO0	nh]
[	SI 45] ONN	7[5	nh]
4	SI 45] SSN	40	r)g]
S5	SI 45] 750	N4	nh]
7N	SI 45] O[I	S]N	nh]
7I	SI 45] OI N	N4	r)g]
rRr HEaRrC v nHRaRT Pae eEH			
Sl	SI 45] 7]5	IS	r)g]
Nl	SI 45] [77	SN0	nh]
Il	SI 45] [SS	OS	r)g]
[l	SI 45] [04	55	rtrynh]
SSl	SI 45] 740	O4	r)g]
S4l	SI 45] [0I	O0	r)g]
O0	SI 45] [0O	O5	r)g]
REsnHRaRT P rRr HEa			
e fo tmto z	SI 45] 77I	[5	rtrynh]
1	SI 45] 4NS	5S	rtrynh]

yeFoi 2Fu 2Huy kyevgd FuF Fi , F y HyFeuH2Fo y Fo, B 23Fy 2F, aHH F, ud F 6VWn , o2D eSuoi A H, H1 e F 2H2 F H, 2u HHF3eFud1 F, 2 , H, Wyeva, o 2Fy 2 2 FF, ud 1 F, 2d y FHHd 2 uFyl HH d, vL v o 2Fy 2 2 LS Ai A H, d y , H2eHyyo lHFv

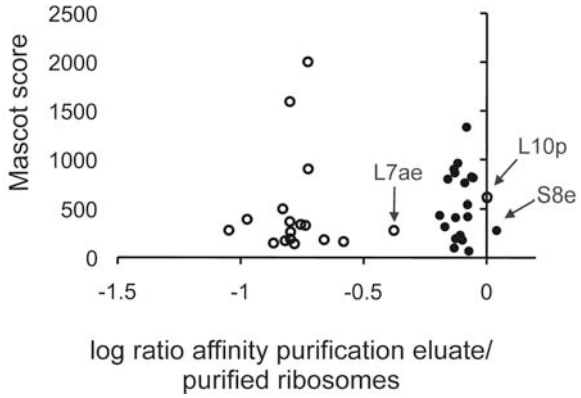


Figure 2: Relative enrichment of individual ribosomal proteins in the MBF1-affinity purification experiment from *S. solfataricus* strain P2 cell lysate. 50% of the ^{15}N-labeled eluate from an MBF1-affinity purification experiment was mixed with 4 pmol of purified ^{14}N-labeled ribosomes. Proteins of the 30S and 50S ribosomal subunits are represented by closed and open circles, respectively.

of the experiment with a standard deviation of ± 0.10 for the normalized ratios.

In order to investigate further if specific ribosomal proteins were enriched during the MBF1-affinity purification, we compared the level of individual ribosomal proteins in the MBF1-affinity purification to intact purified *S. solfataricus* ribosomes. 50% of the affinity purification eluate was mixed with 4 pmol of ribosomes and analyzed by mass spectrometry (Fig. 2). Overall, the results suggested a clear preference for the 30S ribosomal subunit over the 50S ribosomal subunit. The relative amounts of ribosomal proteins for each individual ribosomal subunit were similar for the affinity-purified material and the purified ribosomes with the three exceptions. S8e, L10p, and L7ae were significantly more enriched during the affinity purification. L7ae is not only a component of the large ribosomal subunit, but also part of the RNA-guided rRNA modification machinery [31]. The data suggests that in the experiment intact 30S ribosomal subunits directly interacted with the immobilized MBF1. There was no evidence that additional extrinsic factor are required to mediate this interaction. Intact 50S ribosomal subunits may have also bound to the immobilized MBF1, but alternatively the 50S ribosomal subunits may have been co-isolated with 30S ribosomal subunits in the form of 70S ribosomes.

B n aaHRE E PRePnR P1c RE R ti

e slf nmMr1n 1mtT be h r ifMh MT f rb hftd be le mnmtT
d f b fo M a S Mo oirt b bMr) S d nm
slf nm ifMh 11hftd be le mnm ib slf nmMr d nm eMe nb ifMh
sltr rbM1hftd be kM) O A e r tol f bt Frtd r gi rbMMnmtT
f to M rb Spbe slf nmMr1n 1tT Sd nm 11i 1b tt f rh fto
O70 9 [0 rh Sxo h abtmt1MftbMnmMbe sltr rbM1hftd be le m r
50 9 O0 rh Sxo h abtmt1MftbMnmMbe nbMr fa hftd be le m)

i ftm rnMa hf Mrb rbfMh bMr tT 1 1amb nd nml fTfo
bt lftnM Tfbe f nM r Tf r Mb f bMr tT r th rtim Sd Me
fMtmto 1 nmi irMm) o oirt b bMr tT Sf n 1 be b be lftbM
o Mf b f 1bM 1a Tf MbtbeMhf Mrb tol f bt to tnb abtmt1MftbMm)
eMmihh nmmbe b r th rtim Sf nmtMb md Me eMe o t1 i 1f d Meb
tol 1s m ib Mmtt Mb Tfto be o if Mh mi ftm rnMa rbfMh bMr
kM) O A

]0 fMtmto mT f Mnb 1 r Mmt Mb Mbt O0
r 1 0 fMtmto 1 nmi irMm if Mh mi ftm rnMa hf Mrb rbfMh bMr)
e a r r tr 1a f b M a ftthf o o Mh 11amb t f f r m1bMr M r r
bf r m1bMr mm a r nb MM be]0 fMtmto m a e o M1
ftmm1MFMh.S[H r f be m tr MMr m be MMhfMi bMr tT S Mbe
hf Mrb e r h f nbM1 1a)r 1a o Mtf Tf bMr tT Tf Sd nmTtir p
de f m o tnb S ftbM tyo Mf b d Me]0 fMtmto m r Tf O0
fMtmto 1 nmi irMmk M) O A MMff mi 1mnd f t b M Ttf MT f rb
b ol 1b mp M M bMh be b Sy O0 fMtmto 1 nmi irM Mb f bMr M rtb
1M M t h MMto m j m tr be 7[0 mtflbMr ftCM be o tirbtT
O0 fMtmto 1 nmi irM d nm mbM b bt llftsM b 1a O0 lot1) e
o tirbtT r th rtim Sf mrbM be f r m1bMr nm amd nm
mbM b t Tf b1 nb ef bM m td f)

M M f bt r th rtim Sp f to M rb S tyo Mf b d Me
]0 fMtmto m r Tf O0 fMtmto 1 nmi irMd e r 1 0 lot 1tT f to M rb
Sd nm mill 1 o rb bt f f 1 1amb be b d nm lfthf o o Tf
bf r m1bMr k M) N A o oirt b bMr tT f to M rb Sh n nMr 1m b
tib 70 F tff mtr Mh t be Sb o trto f r b tib N0 F) e Mm
M b tmb M 1a i bt f b M o i 1bM fM bMr tT S if Mh y p
le rto rtr be b M T b r t mfn Tf lftbMnmTfto be fotleMM b Mft m)
tb 1ap S M fm ll f bt b otf M oirtf bM be r be S
o trto fm nm M nM rb Tfto tol fMtr tT be r Mb rnMMmtT

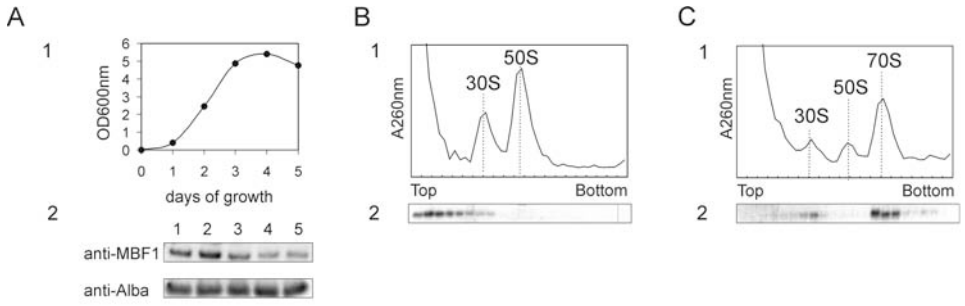
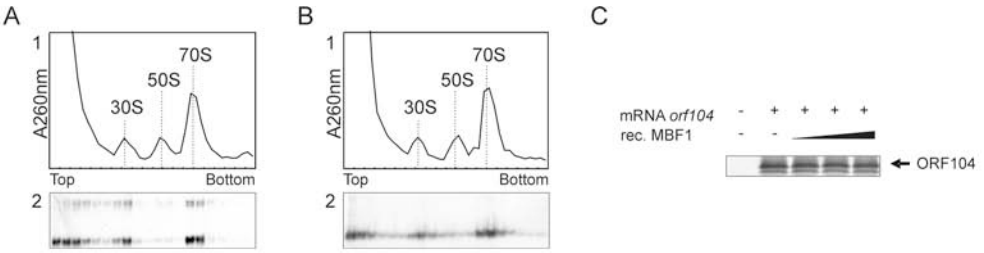


Figure 3: Expression of endogenous MBF1 during different growth phases of *S. solfataricus* strain P2 and its localization in cell lysates after sucrose density gradient centrifugation. (A) Expression of endogenous MBF1 during different growth phases as detected by immunodetection. The upper panel shows the representative growth curve of a *S. solfataricus* culture. The lower pannel shows the immunodetection of MBF1 and the abundant nucleic acid binding protein Alba as control. (B) and (C) Localization of endogenous MBF1 in cell lysate after 70 °C incubation (B) and cell lysate programmed for translation and stabilized by formaldehyde crosslinking (C). The upper panel shows the absorption at 260nm of the fractions after sucrose density gradient centrifugation to identify the position of the ribosomal subunits. The lower panel shows the immunodetection of MBF1

immunodetection and Coomassie-staining (Fig. 1A). Deletion of the N-terminal Zn-ribbon domain of MBF1 (MBF1-C) did not affect the ribosome interaction (Fig. 4B), suggesting that the conserved HTH-domain might be sufficient to mediate the interaction with the 30S ribosomal subunit. In experiments carried out under the same conditions with MBF1-N, the protein was not detectable, probably due to degradation during the high temperature *in vitro* translation reaction. No inhibitory effect on ORF104 synthesis in *in vitro* translation assays was observed after supplementation of up to 300 nM recombinant MBF1 to cell lysate, which corresponds to a more than three times increase compared to the amounts of endogenous MBF1 being present in the assay (Fig. 4C).

Reconstituted complexes

To determine whether MBF1 directly interacts with the small ribosomal subunit, we purified ribosomes at different salt concentrations. Levels of endogenous MBF1 were below the detection limit for ribosomes purified under low salt (100 mM NH_4Cl) and high salt (500 mM NH_4Cl) conditions. Reconstituted complexes consisting of purified 30S ribosomal subunit and recombinant MBF1 could be established with both low salt and high salt washed ribosomes (Fig. 5AB). There was no clear separation between bound



H e G H k A H, H a F H1 o, 2 6 6 6 V a, F H D y 2 F y e u H B 1 F a H u 2 y A H, v K g 6 g K g o 2 y A H, y y D y d o 2 F D y 2 F a u H 1 2 F V 7 y 2 , d F u F y S e e h F 1 F, 2 F d o 2 N5 e 1 H u F H1 o, 2 6 6 K g 6 g V K g v K 6 g l F e u H i h F H a 2 F y H u e A H, 2 O R5, 1 d y 1 F y S u F 2 H d F, A a D 2 F e H y o A H, H a 2 F u o H y H 1 h y S, o 2 y v K O g 1 S, H F 2 F A H, H a F H1 o, 2 6 6 V K g b F 2 H a o, u F y o, B H S, 2 y H a F H1 o, 2 6 6 K 5 f f 5 6 5 g H, F H V F F 2 y A H, v b 2 y A H, d y u o F H S 2 a, 2 F e u F y F, F H a N V F 2 d H, o, F 6 e u H S A H, H a 6 5 U d y F 2 F 2 F D 8 2 H u 2 d H B u e l D F u u y V u o, F V V l F i u y 2 y F y l H d y 6 b 2 y A H, u F A H, d o 2 H S 2 1 2 F 1 e 2 F v

r f l S t r b e h f M r b 2 m i h h m M h b e b o t m b S y O 0 f M t m t o 1 2 m i i r M t o l 1 s m M m t M b i f M h b e r b f M h b M r) o l 1 m d f M i b h b N 2 t f t o l 1 s t f o b M r p M i b h b M r m b e M e f b o l f b i f M t r b t T b M i M h t T S p M l 1 a M h b e b r t b o l f b i f y M i t r T f o b M r 1 e r h m d f f g i M t f S M i M h k b r t b e t d r A r t r b f m b o t b e f m i 1 m t b M t f i 1 y 1 r h b e S p M i M h s l f M r b m f f M t i b d M e S y r S y M r t b a M 1 m b 1 t o l 1 s m d M e b e O 0 f M t m t o 1 m i i r M r b f b e s l f M r b 1 t r M M r m t b e t o M m d f f g i M b t T f o t o l 1 s m b e b d f m b 1 r t i h e b t m i f n M m i f t m r m M a h f M r b r b f M h b M r) S y r b e O 0 f M t m t o 1 m i i r M b M e b e T f o o t f b f r m M r b M b f b M r b e b t i 1 b o t M b e b f r m 1 b M r d M e 1 1 a m b m i t h e e o M 1 f t m m 1 M F M h)

i t h e M m b M M a T S y O 0 f M t m t o 1 m i i r M t o l 1 s m b e f m i 1 b T f t o b e m i f t m r m M a h f M r b r b f M h b M r d m 1 M 1 a b t i r f m b M b e C M M a t T S t f b e O 0 f M t m t o 1 m i i r M b e f T f f M t i b f M t m t o 1 1 b M h m a t b o f o M b e M m t M b M r t r m b r b t T S y O 0 f M t m t o 1 m i i r M t o l 1 s m M t m t o l 1 b M h m a 1 t d T f 1 M t m l f b M r t T f M t m t o y t i r r r t i r S e y n i t f b e M b f b M r d m m b M b t 1 d S 0 0 r k M) A

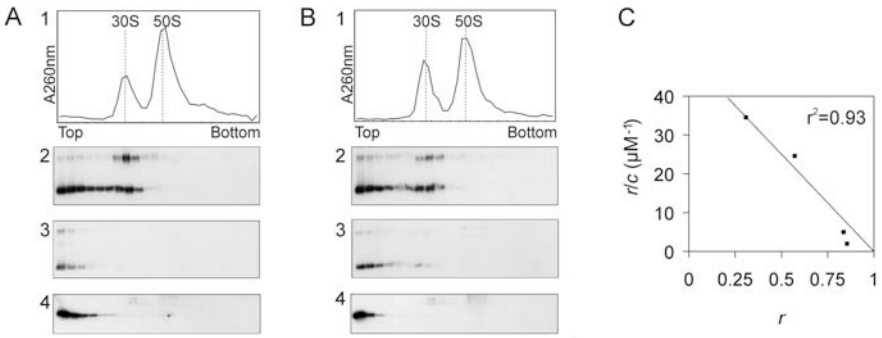


Figure 5: Interaction of MBF1, MBF1-C, and MBF1-N with purified ribosomal subunits. (A) and (B) 100 pmol of ribosomes purified at low salt (A) or high salt (B) conditions were incubated with 100 pmol of recombinant MBF1 or mutant proteins. Upper panel (1) shows a representative A260 profile with the position of the ribosomal subunits. Lower panels show the immunodetection of binding assays for MBF1 (2), MBF1-C (3), and MBF1-N (4). (C) Ribosome pelleting assay to test for co-pelleting of MBF1 with 30S ribosomal subunits. The amount of bound MBF1 was determined by immunodetection of the supernatant. Data were plotted as Scatchard plot assuming stable dimerization of MBF1 where *r* is the fraction of occupied binding sites on MBF1 dimers and *c* is the concentration of free 30S ribosomal subunit.

NMR studies of MBF1-30S ribosomal subunit interaction

In order to study in more detail how the two domains of MBF1 contributed to the interaction with the 30S ribosomal subunit, nuclear magnetic resonance spectra were obtained. The protein gave rise to a well resolved 2D HSQC spectra. The recombinant *S. solfataricus* MBF1 encompasses 173 amino acids including 4 prolines and the C-terminal His_6 tag. 161 cross-peaks were visible in the 2D spectrum (Fig. 6A).

The oligomeric state of MBF1 was characterized by proton T2 relaxation and diffusion experiments. The proton T2 relaxation gave a relaxation rate of 30.46 Hz corresponding to a tumbling time of 8.6 ns. Based on a sphere model for the Stokes equation and a 2 Å hydratation layer, this corresponds to a 24 kDa protein. The diffusion experiment showed homogeneity of the protein preparation and the predicted molecular weight for the diffusion is 15-16 kDa (making the same assumption as for the relaxation concerning the molecular weight). These values would be consistent with a monomeric state of the protein and were in contrast to those obtained from size exclusion chromatography (Fig. 1B).

Purified 30S ribosomal subunits gave rise to a signal that was mainly unfolded, comparable with the signal of ribosomal protein S1 from the *E. coli* 30S

f Mtmto 1mi irMp etd n fbe ff e 1OOf Mtmto 1mi irM Fmr S
tfbet th O7H tbe bff MT f rbar o Mf hMr t The ff e 1OOf Mtmto 1
mi irM o inb tirb Tf be lftbtr nMr 1e OOf Mtmto 1mi irM
metd etoth rtinf Mi nMr d Me Mi nMr t TM Mrbt T Nws S0 ySS o 7n S
tff mtr Mh t be sl b mM t The OOf Mtmto 1mi irM)
y nh bf t f S M mr r r l f mr t T ir 1 1 OO
fMtmto 1mi irM f n 1 m 1 bM ft r Mh t T O4 l Fm M f mr t T
be OOf Mtmto 1mi irM k M)[A e Mi nMr t The n o l 1 tr bM Mh
S r OOf Mtmto 1mi irM M S(S o t 1f f bM k4 Z A se M M
e b fth r Ma d Me M be n o l 1 d Me 1 nb r n td Mi nMr h o t 1 i 1 tr be
tf f t T Nws S0 yS0 o 7n S r Nws S0 ySS o 7n S pf m bM 1a) e M M M b be b be
M Mh t T S t the OOf Mtmto 1mi irM d n f r nMr b)
r t f f t h bM nMe bM bt d e Me to M t T S d n o M 1a tr bf Mi bM h t
be M b f bM r d Me be OOf Mtmto 1mi irM nh bf d f 1nt
f t f T f be M tu 1 b yb fo M 1 to M Sy)25 I ftmmyl Fm
d f t mfn M be 7 nh bfio t T Sy k M)[A) t o l f Mtr bt the
nh bfio t T 1y1 rh be S f n 1 l fT b n f 1a t T be f Fm t b M
T f Sy) r b f nbMh 1ap 11 be t m l Fm f to S be b met d m 1 bM
ft r Mh M be f mr r t TOOf Mtmto 1mi irM nt i 1 nMr t the
yb fo M 1 to M mi hh nbMh be b be M b f bM r d Me be OOf
fMtmto 1mi irM M o M b a be to M k M)[A)

rR tiREarn EaPc RE PHF

r a nhp l t Mb o i b bMr M be g mi r n d 11 n 1 bMr t T be
rbM h r e n r netd r bt 1b f be m r nMM Ma t T be nhf M nbt
o Mth 1a t nM rbM Mh Mhn M i Mh l f t o t o a M k r f Mf p M ntr b 1)
700SA e M M bMh be b S o Me b M b 1a t f M M b 1a Mb f b d Me be
b f r m 1bMr o e M fa M a nb e f e e 1f f r m 1bMr l f bi nMMh r f 11a
f be f M m r nMM bt r b M Mh Mhn OOp ONH r be hf t d be t r NMr m f f
kl O)Op4-0 2 A f M 1a t T b be nh MMa t T r bM Mh Mh) r
tf f bt Mn nbM b d e be f S 1b f m be T b t T l f t o t o a M t r
b f r m 1bMr M
o Mf Mh M 1y f b f r m 1bMr h a mb o m a l f t o t o a M
.7[H r d t 1a if Ma 1M M k t 1ak AA f t h f o o f r m 1bMr h m amp be
f b t T 1i M l f le ra 1 1r M M t f l t f b d n l l f t s M b 1a 0)00N t f
be d M hal nh f M 7 M mr r t T r bM Mh Mp M ht f hf o r b d Me be
li 1Me n 1i t T O)00O 7[H M r t T f o t o a M 1 bt M f m

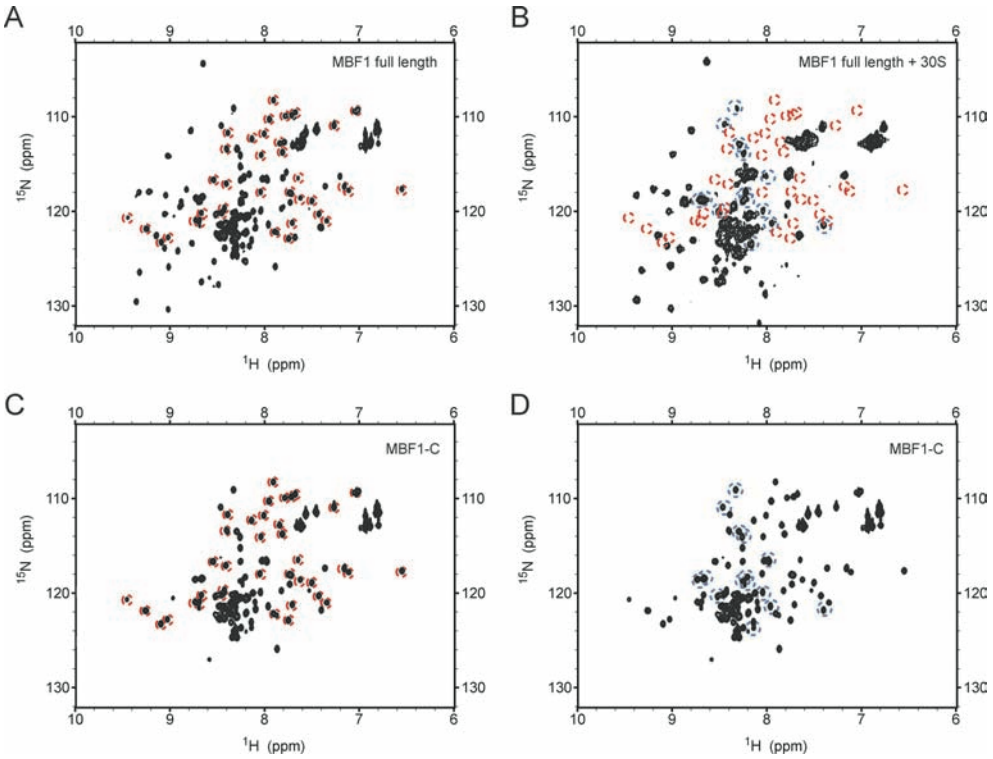
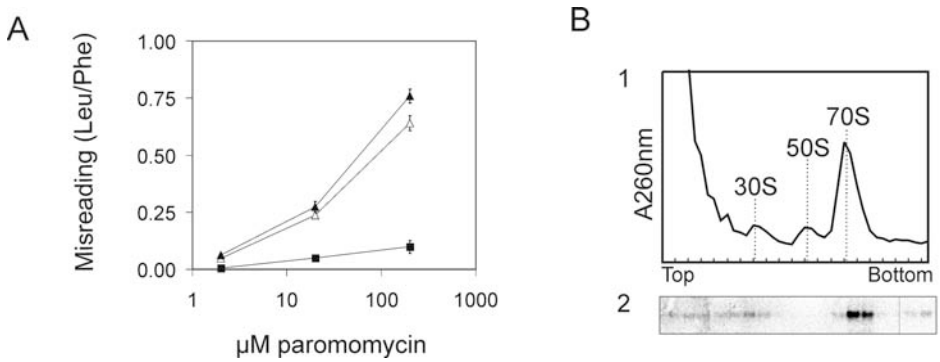


Figure 6: ^{1}H-^{15}N HSQC of MBF1 in absence and presence of 30S ribosomal subunits. (A) Full-length MBF1. (B) Full-length MBF1 and 30S ribosomal subunits in a 1:1 molar ratio (8 µM). (C) & (D) MBF1-C. The cross-peaks marked with open red circles were broadened in the presence of 30S ribosomal subunits, they could all be assigned to the C -terminal domain of MBF1. The cross-peaks marked with blue circles could be assigned to the C-terminal domain of MBF1 but were not broadened in the presence of 30S ribosomal subunits. The cross-peaks from the C-terminal domain of the full-length MBF1 protein and from the isolated C-terminal domain (MBF1-C) exactly overlay.

misreading (Fig. 7A). Both the Δ*mbf1* strain and its parental strain PBL2025 exhibited misreading rates of approximately 0.020 incorporated leucine per phenylalanine in absence of antibiotic. Paromomycin significantly increased the rate of misreading in both strains, but no significant difference in paromomycin induced misreading was found between the parental strain PBL2025 and the Δ*mbf1* strain (Fig. 7A). The about 5-times higher misreading rates observed misreading rates in PBL2025 and the Δ*mbf1* strain might be due to the deletion of about 50 genes in the parental PBL2025 strain [21]. Alternatively, it could be a phenotype of the *S. solfataricus* 98/2 strain from which PBL2025 was derived.



Hen b F 2Ha 6H, e 1 H1 D o, o S F 1 oyuF o, B o, eHhDK gW uF 2F 2u yh AH, y yDyv K g o 1 oyuF o, B u2F y F2F D2 F o, Hue HuA H, Ha hFS o, F o, 2H eHhDel F, Dh o, F D eHhDK gW uHBu 1 F u o HyH1 Fyv o FuF, 2 FHhDy2Fyd FuF2Fy2F -V 7 KHeF, 2u o BhFg e F, 2h y2u , OSON K hHyF 2u o BhFg d o 2DeF y2u , O K hHyF yLS Fg uHu y uFeuFyF, 2H, F y2 , 7 FjoA H, vK g b F2Ha655x e 1 H1 D o, H, 2l F HW oBuA H, HaF, HBF, HSy 6 d o l f5 u o HyH1 h yS S, o2y p5 u o HyH1 Fy o FhDy2F euHBu 1 F Hu2u yh A H, v

t s i be l t mM MM a be b l f t o t o a M t i b t o l b S T f O0 f M t m t o 1 h i r M M h d b f o M be 1 t 1M b M r t T S M 1 1 1am b l f t h f o o T f b f r m 1 b M r r mill 1 o r b d M e S 0 0 Z l f t o t o a M) o o i r t b b M r f n 1 be b S d m m M 1 o t m b a m t M b d M e] 0 f M t m t o m k M)] A)

oy Syyd H,

he t d e f be b be S t f b e t 1 t h f t o M b f b m d M e be O0 f M t m t o 1 i r M) e TCM M a l i f M M b M r sl f M r b r be f t r m b M i b S y O0 f M t m t o 1 i r M t o l 1 s m M l 1 M b be b be M b f b M r M M M b) i f be f o t f p be t y f b M r b M r t T r t h r t i m r f t o M r b S d M e h f r m 1 b M h] 0 f M t m t o m M M 1 am b t f f t t f b m be l ean M 1 t h M 1 f 1 n r t T be M M b f b M r) e M m t M b M r t r m b r b m b f o M a M T f r b o b e t m T f be S y O0 f M t m t o 1 i r M t o l 1 s k 3 7 O r A f M M m M f f r h m f T f M m b r be M m t M b M r t r m b r b T f be b f r m 1 b M r 1 f h i 1 b t f [M M h b t be l 0 f M t m t o 1 i r M . 7 I H r be a f d 1 1 1 d be l ean M 1 t h M 1 t r r b f b M r t T f M t m t o m) e M m i h h m m be b be o R f f b M r t T S d M m t M b d M e f M t m t o m M) e m M M a t T he M M b f b M r M f be f t r O M o a be b e b be b r t M b M M h b t l 0 f M t m t o 1 i r M m d n t m f n)

Neither the presence of synthetic RNA poly(U) nor tRNA appeared to affect the stability of reconstituted MBF1-30S ribosomal subunit complexes (data not shown).

Although our data suggest a function for the archaeal MBF1 ortholog in the translation process, its exact role therein remains unclear. No inhibitory effect on *in vitro* translation was found for recombinant MBF1. This might indicate that MBF1 does not inhibit translation initiation complex formation, which is the rate-limiting step of translation under normal conditions. Either because MBF1 binding to the 30S ribosomal subunit is compatible with the translation initiation complex formation or MBF1 associates at a later stage during the translation cycle with the 70S ribosomes and still remains bound to the 30S ribosomal subunit after subunit dissociation.

Because MBF1 is present only in substoichiometric amounts compared to the ribosome, it might be present only on a sub-fraction of the ribosomes engaged in translation. However, the fraction of ribosomes interacting with MBF1 during translation could be higher due to rapid association and dissociation of MBF1. Indeed, the dissociation of MBF1-30S ribosomal subunit complexes during sucrose density gradient centrifugation as well as the dynamic of the interaction as detected by NMR suggests such a highly transient interaction. This could explain why eukaryotic MBF1 has not been found to be associated with the ribosome so far. While our experiments point to highly dynamic MBF1-30S ribosomal subunit complexes, this is not necessarily true for an MBF1-70S ribosome complex that the co-migration patterns of MBF1 suggest to exist. Unfortunately, the instability of crenarchaeal 70S ribosomes prevented the reconstitution of MBF1-70S ribosome complexes and their characterization.

The MBF1-30S ribosomal subunit interaction is mainly mediated by the HTH-domain of the archaeal MBF1 ortholog. Using GST-pulldown assays with purified recombinant proteins, an interaction between the C-terminal HTH of yeast MBF1 and the general transcription factor TBP could be demonstrated [6, 7]. We have not been able to detect any interaction of *S. solfataricus* MBF1 with TBP in the affinity purification experiment using either cell lysates (Table 1) or purified, recombinant TBP (data not shown). The interaction between TBP and MBF1 might therefore have evolved in the eukaryotic domain only, although co-evolution of TBP and MBF1 in Archaea has been suggested [7].

M i o nb r bM1 nM r mi hh nbnbe b M a nb S t i 1 e n nt T f r f t hr M 1 t 1 M bf r m1 bM r nd 1 p r 1 be b M o M eb M b f bd M e be N O no 1 1 f M tnto 1 mi ir M M nbp o i b bM r r 1 bM r t T 1 bf mr nM M M a t T he nbf M nbt o M th1a tnM 1 r bM M bM nbe b bf h b be N O f M tnto 1 mi ir M p M i M h1 f to to a M S 0 H t r p a nb S M M illf nntf t T f M tnto 1 6 S f o ne M bM h S 0 yS 7 H e M p nf r t f lft b M yl ft b M M b f bM r n M a nb mi hh nbn M b f bM r nt T S d M e mn f 1 lft b M nbM nt 1n M bf r m1 bM r M M M bM r (bf r m1 bM r M M M bM r t f n y7 p y7 p r y[be e 1M m S p be bf r m1 bM r f hi 1 bt f S p r be 1 y nnt M b lft b M 7 0 OI H

e f e ynh M M r ryf M tr t to M t T S m o n bt 1 f bM M b t r 1 a d F 1 a M be M b f bM r d M e O O f M tnto 1 mi ir M M nM r b f to be nh bf) r ryf M tr t to M nm h r f 1 a b M b M r M b f bM r n r ryf M tr t to M t T f e 1 S t f be t 1 th n b M e b be i n M b f bd M e r t be f 1 M r d e M be hM i 1 b r tim a M nt be O O f M tnto 1 mi ir M e M 1 M r t i 1 be I O f M tnto 1 mi ir M nf i bi f 1 f t f lft b M be b n l be M r bM M bM r M be Sy TC M M a lif M M bM r sl f M rb

e M bf r m f M bM r r bf r m1 bM r f nh bM 1 a m l f b lft nnn M i F fat b mp be a f t i 1 M f e O[H nh r f M sl 1 M M h be t mfn nnt M bM r nd M e be bf r m f M bM r r bf r m1 bM r o e M f M n M i F fat b m r f e dti 1 ir bM r M h t T S M be bf r m1 bM r M M M bM r tr r d 1 a bf r m f M o) r mi e nm r f M S dti 1 M f m be 1 t 1 tr r b f bM r t T O O f M tnto 1 mi ir M nf b be nM t T bf r m f M bM r M M M bM r r ti 1 M f m be f a be TC M M r a t T f fi M o r b T he O O f M tnto 1 mi ir M t he o) nM h he M e alt be nM dti 1 f gi M til 1 bf r m f M bM r xbf r m1 bM r na m o be b M iff r b1a rt b n M 1 lt f td n f t if b lft n M r t n M r t f nh M M M bM r t T S d M e be bf r m1 bM r M M M bM r to l 1 s m r b lf th f nn M be nn o 1a t T f e 1 bf r m1 bM r M M M bM r to l 1 s m d M 1 1 td t b nb t f r M b f bM r t T be f e 1 S t f be t 1 h d M e be bf r m1 bM r M M M bM r to l 1 s M he r f a T b if O J H

Acknowledgements

The Anti-Alba antiserum was a kind gift from Malcolm White (St. Andrews University). FB would like to thank Fred van der End (Wageningen University) for assistance with the growth of *S. solfataricus* cells. FB was supported by an EMBO short term fellowship (ASTF304-2008). This work was financially supported by a VICI-grant to JO by the Netherlands Organization for scientific research (grant number 865.05.001).

References

1. Aravind L, Anantharaman V, Balaji S, Babu MM, Iyer LM: **The many faces of the helix-turn-helix domain: transcription regulation and beyond.** *FEMS Microbiol Rev* 2005, **29:**231-262.
2. Aravind L, Koonin EV: **DNA-binding proteins and evolution of transcription regulation in the archaea.** *Nucleic Acids Res* 1999, **27:**4658-4670.
3. Blombach F, Makarova KS, Marrero J, Siebers B, Koonin EV, van der Oost J: **Identification of an ortholog of the eukaryotic RNA polymerase III subunit RPC34 in Crenarchaeota and Thaumarchaeota suggests specialization of RNA polymerases for coding and non-coding RNAs in Archaea.** *Biol Direct* 2009, **4:**39.
4. Werner F, Weinzierl RO: **Direct modulation of RNA polymerase core functions by basal transcription factors.** *Mol Cell Biol* 2005, **25:**8344-8355.
5. Grunberg S, Bartlett MS, Naji S, Thomm M: **Transcription factor E is a part of transcription elongation complexes.** *J Biol Chem* 2007, **282:**35482-35490.
6. Takemaru K, Harashima S, Ueda H, Hirose S: **Yeast coactivator MBF1 mediates GCN4-dependent transcriptional activation.** *Mol Cell Biol* 1998, **18:**4971-4976.
7. Liu QX, Nakashima-Kamimura N, Ikeo K, Hirose S, Gojobori T: **Compensatory change of interacting amino acids in the coevolution of transcriptional coactivator MBF1 and TATA-box-binding protein.** *Mol Biol Evol* 2007, **24:**1458-1463.
8. Li FQ, Ueda H, Hirose S: **Mediators of activation of fushi tarazu gene transcription by BmFTZ-F1.** *Mol Cell Biol* 1994, **14:**3013-3021.
9. Kabe Y, Goto M, Shima D, Imai T, Wada T, Morohashi K, Shirakawa M, Hirose S, Handa H: **The role of human MBF1 as a transcriptional coactivator.** *J Biol Chem* 1999, **274:**34196-34202.
10. Hendrick JL, Wilson PG, Edelman, II, Sandbaken MG, Ursic D, Culbertson MR: **Yeast frameshift suppressor mutations in the genes coding for transcription factor Mbf1p and ribosomal protein S3: evidence for autoregulation of S3 synthesis.** *Genetics* 2001, **157:**1141-1158.
11. Costanzo MC, Mueller PP, Strick CA, Fox TD: **Primary structure of wild-type and mutant alleles of the PET494 gene of Saccharomyces cerevisiae.** *Mol Gen Genet* 1986, **202:**294-301.
12. Culbertson MR, Gaber RF, Cummins CM: **Frameshift suppression in Saccharomyces cerevisiae. V. Isolation and genetic properties of nongroup-specific suppressors.** *Genetics* 1982, **102:**361-378.

6f v F H - r C P r E a r H c RE P e P c RE s l r H c RE RE r r r H P r C T t S
O55 N q ti fi b O R O W O R N v

6U v l F q o, B l q H, a H, d u q o j, Hj o q H a q d D F k q l , V W F d F u q q H S yF,
q S u A y q F H HHu y q F 2 h - C r R T P r r E R T R R r C r E r C R E e P R R R e a
a R r r H e a x i .
6N v S l h F u q H s ye F d l q H d h B - r E H c RE E e P R c a 5 e E R R r C r E a
E R H E r C r r a e e E H a g E E R R r C r s E E r R F T r a R R r C
r C r r H e T e P R R R e a H R r r H e a i s S S S t i q fi, bUN6pWNf Uv
6R v F, F H q H, F d - E d H r R a r e H a R r r C P r E a P c RE P H E H t c RE i . S of
O55 p q G fi v b p t i W5 t i v
6p v H, H q d 1 S H, d F, F H d S B B F u H q H, F d - H a s c E a H E P a R E r r R P H E
r E a P c RE P H E H t c RE H E r C r C n T R C H r n C R E e P R R R e a a R r r H e a i
t i t i q fi G f v b p t i W5 t i v
6I v l , F q d d 2 F a H r d F u d q F, F H d q d k q F H, F d - e E c RE P
E P F a H a R r C r n E a P c RE r r R r x d - H E r C r C n T R C H r n C R E e P R R R e a
a R r r H e a i S O55 p q l - b p 55 W 6 f v
6t v d H q H, o e F H - P e P c RE R r n r R r H E B c E c RE R H E r a n R T T H E R H E
a 5 e E r r i t i q t i N x b f 6 t i W O R v
O5 v B F u q H, H d - n H H E s a R H e T r R F F a e P r s R F n F P T H P P
P r n R C R n a H R n r C a r n c RE R r n r R r H E a H E r C n E n R T t i a R t i v v A i
t i t i l p q fi l l b f R I W p t i v
O6 v l F H u 2 q o 3 d q H, B q d 1 q , q u H k q l S 1 - e n E E E E
C n r r n H c RE R E R T n e n F n a H a r E H E r C C F n r C n T R C H r n C R E e P R R R e a
a R r r H e a F a e a R R r E H a n e c RE i . O55 U q t i N l b U O p W f p v
OO v e 2D q y y F u q F u A B q a F u c 1 e q H S u d q , o q q F d 1 , q
l u F o F u q d F u H o, y c o q F 2 h - p R r a r E n r a p R r n r C r C n T R H R C H r n C R E
e P R R R e a a R R r r H e a i / . O55 t i v
Of v , o F u y q q i 2 l F u q F 2 F u q o , 1 , q F H y q , F u F u c F, u S, y q ,
F u H y 2 q u B l 2 - R E a r n e c RE R E r n P r n R E T r R R H a T H e P R R R e a
a R r r H e a e a H E r M R s H T E a H R E P P P r n R C R n a H T g a r P H a R r R H P H E
E T H n R r n F E F a H a i . S O55 R q l b 6 N 6 I W N O t i v
OU v B S D F, q H H u l - n R r H s n R r H E H E r n c RE a a F a b P H T H E c E P a R a k t l
H E r n c RE a i . S O55 R q fi b 6 f N W f t i v
ON v F, F H d q k q F, F H, k q F, q H, F d - e E c RE E E H n H R a R T P
H P H 1 c RE R P E l g r n E a P c RE P e P r R n a C n F F n C E E e A n F i s S S S O55 t i q fi, b O N R W O R p v
OR v H, F q h 2 S u q, k q 1 d y - T H E R F R a H s H E e T H a r n E a P c RE H E
r C n T R C H r n C r n H i . t i t i q x t i G b U l W U v
Op v F H, o, B q H H 1 l S q u H S, y g , F u H y 2 - R P E R T e R n R r H E n H E
E r R n a H E r n C b n H H E r C r T H E a t S O55 t i q fi, b N O W p v
Ol v k d q F 1 S c F B q y l d 1 q F o y u H y F q q F u H c F, q , d
- E c c RE R a r C R R n R T H E E r C a a RE n F a r n e r e n R a r C r n E a n H c RE P
R R c l r R n t i S S S S 6 t i t i q G b U 6 N W O U v
Ot v u H B q B , F D q S H q B, 2 l F, c H d S H B, S q s o q c S o y y q F 2
h - P R P P E a R a r H E r C F F a r C n T F a H n l H H i . s
O55 R q G G v b R f p W U f v
f 5 v , - e E c RE E 5 e E c r c l r R r T H a a e a H i . o
, b t i N O W N l v

31. Kuhn JF, Tran EJ, Maxwell ES: **Archaeal ribosomal protein L7 is a functional homolog of the eukaryotic 15.5kD/Snu13p snoRNP core protein.** *Nucleic Acids Res* 2002, **30:**931-941.
32. Lecompte O, Ripp R, Thierry JC, Moras D, Poch O: **Comparative analysis of ribosomal proteins in complete genomes: an example of reductive evolution at the domain scale.** *Nucleic Acids Res* 2002, **30:**5382-5390.
33. Cammarano P, Teichner A, Londei P, Acca M, Nicolaus B, Sanz JL, Amils R: **Insensitivity of archaebacterial ribosomes to protein synthesis inhibitors. Evolutionary implications.** *Embo J* 1985, **4:**811-816.
34. Ruggero D, Londei P: **Differential antibiotic sensitivity determined by the large ribosomal subunit in thermophilic archaea.** *J Bacteriol* 1996, **178:**3396-3398.
35. Tarassov K, Messier V, Landry CR, Radinovic S, Serna Molina MM, Shames I, Malitskaya Y, Vogel J, Bussey H, Michnick SW: **An in vivo map of the yeast protein interactome.** *Science* 2008, **320:**1465-1470.
36. French SL, Santangelo TJ, Beyer AL, Reeve JN: **Transcription and translation are coupled in Archaea.** *Mol Biol Evol* 2007, **24:**893-895.
37. Hasenohrl D, Fabbretti A, Londei P, Gualerzi CO, Blasi U: **Translation initiation complex formation in the crenarchaeon Sulfolobus solfataricus.** *RNA* 2009, **15:**2288-2298.

?SeelF1 F, 2?l D??dF

```
taatacgactcactatagggagaGAGGTGAAAATATAAATGTCTCAAAGCTTTGAGGGAGA
     T7 promoter       SD motif       M  S  Q  S  F  E  G  E

ATTAAAAACAATTCTCAGAAGTGGCAAAGTTATTTTGGGAACAAGGAAGACATTGAAATT
  L  K  T  I  L  R  S  G  K  V  I  L  G  T  R  K  T  L  K  L

ATTAAAGACAGGAAAGGTAAAGGGAGTAGTAGTTTCTTCTACATTAAGGCAGGATCTAAA
  L  K  T  G  K  V  K  G  V  V  V  S  S  T  L  R  Q  D  L  K

AGACGATATAATGACATTTTCAAAATTTTCTGATATTCCAATTTATCTCTATAAAGGTAG
  D  D  I  M  T  F  S  K  F  S  D  I  P  I  Y  L  Y  K  G  S

TGGATATGAATTAGGGACATTATGCGGTAAACCTTTTATGGTATCTGTTATAGGTATAGT
  G  Y  E  L  G  T  L  C  G  K  P  F  M  V  S  V  I  G  I  V

TGATGAAGGGGAATCAAAAATTTTGGAGTTTATTAAAGAGGTGAAGCAATGAGTGCCAGA
  D  E  G  E  S  K  I  L  E  F  I  K  E  V  K  Q  *

AATTAAATTAAC
```

?e **PT ?Er?nF??H?en??n?b??FLSF, ?F?Ha?2l F????b????2u? yh?A H, ?2F1 el?2F?SyF??o, ?2l oy?y2S?D???y F??H, ?2l F????*rc*1 ????T6J?o, ??Hl??f?s Fuy?d o2l ?1 S2?A H, y?o, ?2l F?, S?lFHA?F?? ??euH2Fo, ?yFLSF, ?F?o, ?d??2F??y ??Hl??f?s Fuyv??l F??p?euH1 H2Fu?yFLSF, ?F?SyF??aHu????b????2u? y?uceAH, ?? ???2l F??l o, F??f?n B?d , H?K??g?1 HAa?d F?S, ?Fuho, F?v??D, 2l FA????????Ha?2l oy?yFLSF, ?F?d?y ??lH, F??j?d?????? ????????o, 2H???e?lS Fy?uce2??Fuuj?2F??j F?2Hu?? ???2l F?elfy 1 o?d?y ?ho, F?d dkF?d?l o2l ?????????euoHu2H-????b????2u? y?uceAH, v

CHAPTER 6

Identification of an ortholog of the eukaryotic RNA polymerase III subunit RPC34 in Crenarchaeota and Thaumarchaeota suggests specialization of RNA polymerases for coding and non-coding RNAs in Archaea

Fabian Blombach[1], Kira S. Makarova[2], Jeannette Marrero[3], Bettina Siebers[3], Eugene V. Koonin[2], and
John van der Oost[1]

[1] Laboratory of Microbiology, Wageningen University, Dreijenplein 10, 6703 HB Wageningen, Netherlands
[2] National Center for Biotechnology Information, NLM, National Institutes of Health, Bethesda, Maryland 20894, USA
[3] University Duisburg-Essen, Faculty of Chemistry,Biofilm Centre, Molecular Enzyme Technology and Biochemistry, Lotharstrasse, MF Building (F 166), 47057 Duisburg,Germany

Biology Direct 4:39

□□2Jv□ J

□t □□i □□f□□□□□een □adof□ □□□rd□auif □□t □l an □f i t □fai □□oo t □□of□□□□i N□g of□t □□□ i □□z□□ of t □fy□h r□f N□r □r t f□□□□□fi□un □a□o□□in f□g□o□□□f□□a□□□□ □□f□□□fi □ f□□□fa□t o□a f f i t □i □□of □□w□□□e□oo□o□i □□□i □ t □□i a□t i t N□i □ t □□□□□ok□□a□□□□□ □t □i □□□i t□u□i t □□ □□□□fi□un □a□o□□ f□□f□a□o□n □e□o□f□□□ □r d□auif □□□□□□ fi□un □a□o□□□□l f□□a□of □□f□fi □f□□□or □r t f□□in f i o f i t□□□a□□l □□ □□t f □u□ □a□□□□□e□i af□i ei □o□i □□f□□□□r d□auif □□□□□□fi□un □a□o□□□□□or □r t f□□□□z-□□ □□t i n □□□i t f□gf□□t □□uo o□or f f i afo□□□a□t □f i t □i □□f□ o□□a□□□□□e□f ai f□t □t □f□□□ fa□t o□a f f i t □i □□t i t N□i □ t □□□□□ok□□□□o□□wt □ t □o□or □□□of□f□□f□□a□t □f i t □e□ o□f □a□f i t □l □□□□□□fi□un □a□o□o□□l a□ai f□t N□i □ t □□□□t □o□□t □□i t N□i □ t □□□□o□ n □□f□l a□□□f□□l□□□a □t □l □□□□□r d□auif □oR

□□go□V□v2□□l o2□□ tt□,□□V□ □V□go□V□□□□O□□r□v□o□□ 2J□vK□ □□ □□□□Jv□□1□□SvJ□w□□□r SK o□rJ□□□□O□ □uvdo□□ o r□□. M□tC□□v □h□

Findings

All Eukaryotes possess 3 distinct, multi-subunit RNA polymerases (RNAPs): RNA polymerase I (transcription of 16S and 23S rRNA), RNAP II (transcription of protein-coding mRNAs), and RNAP III (transcription of 5S rRNA, tRNA and some other small non-coding RNAs). Plants have two additional RNAPs involved in the transcription of small interfering RNA [1].

RNAP III has counterparts (either identical or paralogous) to all subunits of RNAP I and RNAP II [2]. In addition, RNAP III possesses the loosely bound RPC82/RPC34/RPC31 sub-complex. This sub-complex is present in all Eukaryotes, although RPC31 is missing in two major eukaryotic lineages (Alveolates and Excavates) [3]. Transcription initiation by RNAP III requires, among others, the TBP and TFIIIB70 proteins. TBP is shared with RNAP II, and the N-terminal region of TFIIIB70 is homologous to the RNAP II factor TFIIB, whereas the C-terminal region is specific for TFIIIB70. The archaeal RNAP (aRNAP) resembles RNAP II in its subunit composition [2]. Furthermore, the aRNAP machinery employs the transcription initiation factors TBP and TFB, which are orthologs and functional counterparts to the eukaryotic TBP and TFIIB/ TFIIIB70, respectively [4].

In the RNAP II and aRNAP machineries, TFIIB and TFB are thought to recruit the RNAP directly to the transcription pre-initiation complex. In contrast, RNAP III requires the RPC34 subunit to mediate the interaction between TFIIIB70 and RNAP III [5-7]. Both the conserved N-terminal region and the unique C-terminal region of TFIIIB70 contribute to RPC34 binding [6, 8]. Given the conservation of RPC34 in all eukaryotes and its central role in the recruitment of RNAP III to the pre-initiation complex, it seems likely that RPC34 played an important role in the evolution of the RNAP III transcription system. To address this possibility, we set out to identify potential archaeal homologs of RPC34.

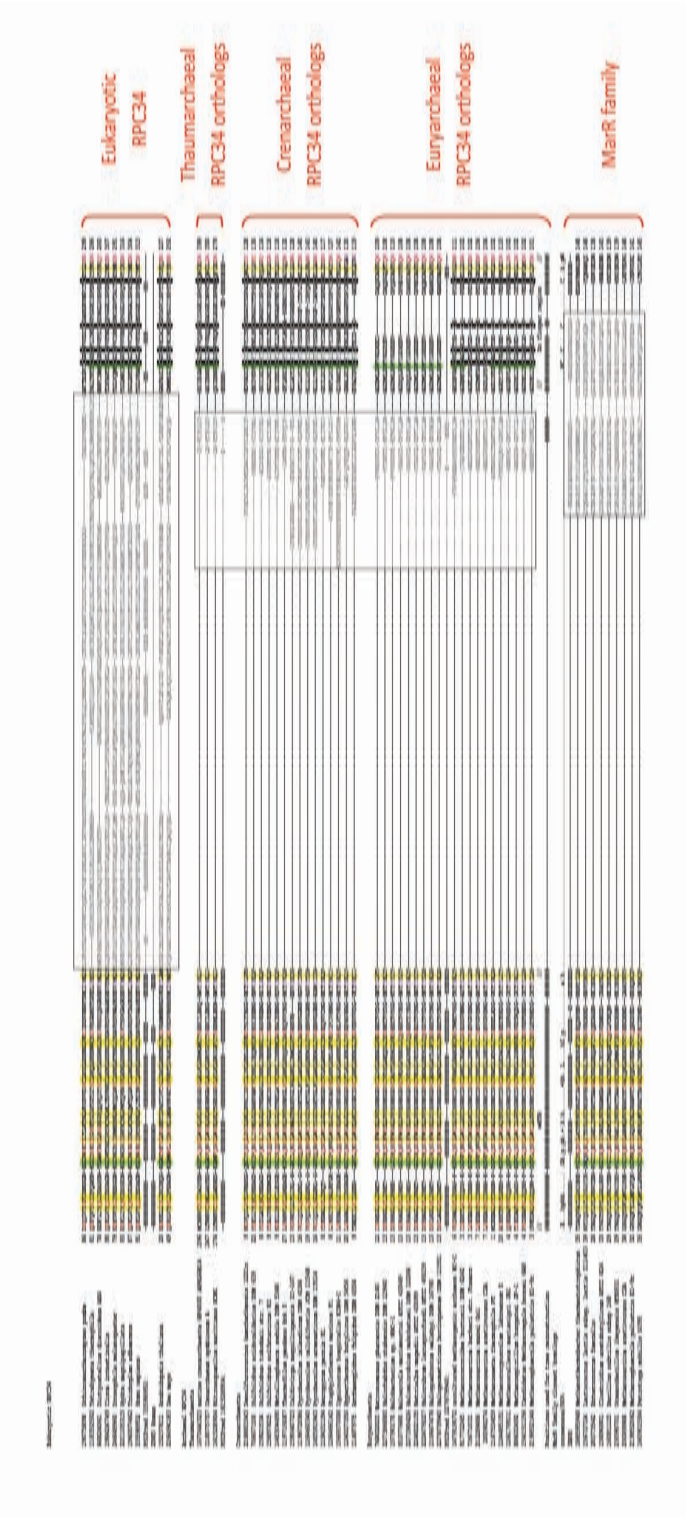
Identification of archaeal homologs of RPC34

Using PSI-BLAST search [9] (against the RefSeq database [10], with default parameters) with human RPC34 as the query (GI: 149640989), we detected a hit to a *Cenarchaeum symbiosum* (strain A) protein [GI: 118575757 with E-value=$5x10^{-5}$] after the first iteration. Reciprocal search starting from the *Cenarchaeum symbiosum* sequence [GI: 118575757; 244-362 aa] identified the first eukaryotic RPC34 homolog [GI:157138209 with E-value=$4x10^{-10}$] after the first iteration. All archaeal orthologs can be retrieved after the first iteration in the

course of the same search (the complete information is available at http://www.biology-direct.com/content/4/1/39). We identified apparent orthologs of RPC34 in all crenarchaeal and thaumarchaeal genomes as well as in several lineages of Euryarchaeota but not in Candidatus *Korarchaeum cryptofilum* OPF8, the only Korarchaeote sequenced so far. None of these archaeal sequences are annotated as RPC34 homologs in the Refseq database. In agreement with the PSI-BLAST results, a Conserved Domain Database search [11] with various crenarchaeal and thaumoarchaeal sequences as queries identifies the statistically significant similarity (E-value ~0.001) of their C-terminal domain to a profile pfam05158, RNA polymerase RPC34 subunit. A similar result was obtained using HHPRED search [12]. For the same *Cenarchaeum symbiosum* A query, pfam05158 (RNA polymerase Rpc34 subunit) was detected with E-value=$6.6x10^{-23}$; in the same HHPRED search, the sequence corresponding to the structure of human RPC34 winged helix-turn-helix (wHTH) domain [PDB:2dk5] was detected with E-value=$2x10^{-11}$. The next most similar family of wHTH-domain-containing proteins was the MarR family of transcriptional regulators (pfam010470, with E-value=$1.2x10^{-9}$). The latter observation is also consistent with the PSI-BLAST search results of the HTH region of archaeal RPC34 orthologs in which MarR family sequences were identified as the closest hits. Most likely, this relationship between RPC34 and MarR is the cause of the misannotation of some of the apparent archaeal orthologs of RPC34 as MarR family transcriptional regulators [e.g. GI:18313992]. Thus, the N-terminal region of archaeal RPC34 orthologs contains a wHTH domain, whereas the C-terminal domain is a distinct Zn-finger domain shared with most eukaryotic RPC34 sequences.

The multiple alignment of the eukaryotic RPC34 sequences and their archaeal orthologs reveals conservation of two regions (Figure 1a). In agreement with the above observations, the first region corresponds to the N-terminal wHTH

[illegible]


Eukaryotic RPC34
Thaumarchaeal RPC34 orthologs
Crenarchaeal RPC34 orthologs
Euryarchaeal RPC34 orthologs
MarR family

domain (with all structural elements of wHTH, namely, three α-helices and two β-strands, preserved) whereas the second conserved region corresponds to the Zn-finger domain with the unique CxxC-x(3-5)-C-x(4-10)-C signature. There are substantial differences in the Zn-finger domain architectures of the euryarchaeal domains, on the one hand, and the crenarchaeal, thaumarchaeal and eukaryotic domains, on the other hand. In particular, the Zn-finger signature cysteines are not conserved in all sequences from Halobacteriales. All eukaryotic sequences contain a structured insert between the wHTH and the Zn-finger domains (according to PSIPRED [13] secondary structure prediction) that probably represents a distinct domain. Thaumoarchaeal proteins contain an extended region of low complexity N-terminal of the wHTH domain.

[illegible]

We constructed a phylogenetic tree from the alignment of the wHTH and Zn-finger domains of the eukaryotic and archaeal RPC34 orthologs, using the MarR family wHTH domain as an outgroup (Figure 1b). Consistent with the apparent synapomorphies in the Zn-finger domain architecture (see above), the phylogenetic analysis shows that eukaryotic proteins group with crenarchaeal and thaumarchaeal sequences with reliable bootstrap support, excluding all euryarchaeal sequences (Figure 1b and Supplementary Figure 1). Moreover, the eukaryotic lineage is rooted deeply within the crenarchaeal-thaumarchaeal subtree (Figure 2), suggesting that eukaryotic RPC34 indeed originates from an ancestor that belonged to this group of archaeal proteins.

[illegible]

To gain insight into possible functions of the archaeal RPC34 orthologs, we analyzed the genomic context of the respective genes. In thaumarchaeal and crenarchaeal genomes, the RPC34 genes co-localize and are predicted to be co-transcribed with several genes for proteins involved in modification or processing of tRNA and rRNA (Figure 1c). In the majority of crenarchaea, the RPC34 gene is also potentially co-transcribed with a gene for a TFB paralog (COG1405). Generally, archaeal genomes encode at least two TFB paralogs, so an intriguing possibility is that the crenarchaeal RPC34 ortholog interacts with a specific TFB paralog analogously to the interaction of eukaryotic RPC34 with the TFIIB paralog TFIIIB70. Genes encoding the euryarchaeal RPC34 orthologs, with the exception of those from Halobacteriales, are predicted to be co-transcribed with genes for Sm-like protein paralogs (COG1958). The *Archaeoglobus fulgidus* homolog Af-

Sm2 has been shown to co-immunoprecipitate with RNase P RNA, and Sm-like proteins are generally believed to form ribonucleoprotein complexes [14].

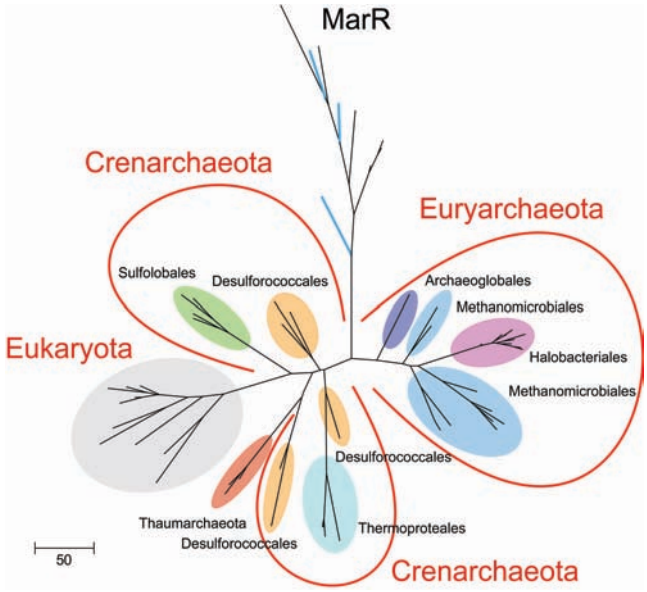
Possible role of the archaeal RPC34 ortholog in transcription

The genomic context of the archaeal RPC34 ortholog, as well as the analogy with the eukaryotic RPC34, suggest that these archaeal proteins might be involved in transcription of rRNA and tRNA genes. It has been shown that, in a reconstituted *in vitro* transcription system from *Sulfolobus shibatae* transcription from rRNA and tRNA promoters could be successfully initiated in the absence of the RPC34 ortholog [15]. Hence, there is no strict RPC34 requirement for recognition of these promoters and recruitment of the aRNAP. Nevertheless, a regulatory role of this protein in the transcription of structural RNAs by aRNAP appears likely. The wHTH motif might mediate protein-DNA-interactions given that the eukaryotic RPC34 was cross-linked to DNA in transcription initiation complexes [16] but, to our knowledge, RPC34 has not been reported to contribute to promoter recognition. Electron microscopy revealed the position of the RPC82/RPC34/RPC31 sub-complex in the core RNAP III close to the "clamp" formed by the N-terminal part of the largest subunit, C1 [17]. The "clamp"-domain is conserved in all multi-subunit RNAPs, but an RNAP III-specific region is thought to be important for RPC34 binding specificity [3]. The archaeal RPC34 ortholog might similarly recruit aRNAP to the transcription pre-initiation complex via the "clamp"-domain and so enhance the transcription of structural RNAs.

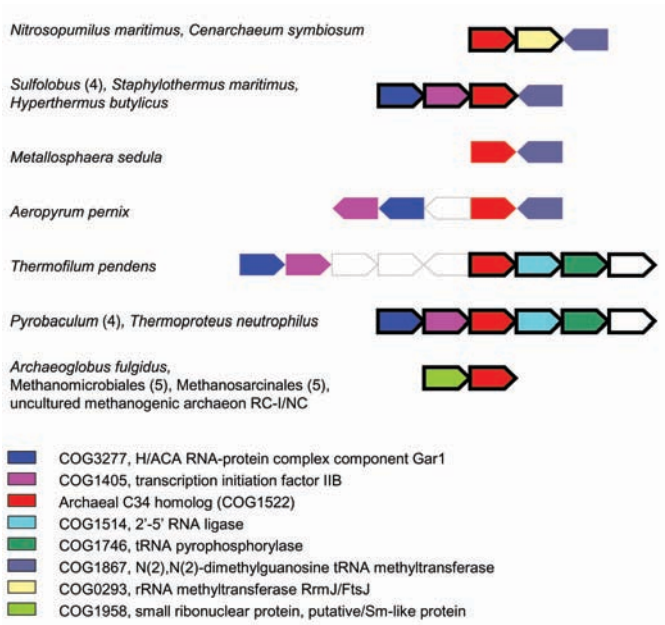
On the origin of eukaryotic RNAP multiplicity

The detection of a RPC34 ortholog in Archaea suggests that the separation of RNA polymerases into dedicated forms for the transcription of protein-coding genes and genes for structural RNAs (eukaryotic RNAP II and RNAP III, respectively) might have evolved already in Archaea and was inherited by Eukaryotes from the "archaeal parent". In this scenario, the archaeal RPC34 ortholog would modulate the specificity of the single aRNAP, whereas in Eukaryotes the specialization deepened as a result of the duplication of the genes coding for other RNAP subunits and general transcription factors. Experimental analysis of the functions of the archaeal RPC34 ortholog will provide a direct test of this hypothesis.

The nature of the archaeal "parent" of eukaryotes is a wide open question [18, 19]. Detailed comparison of individual functional systems allows partial reconstruction of the gene repertoire of this elusive entity. With respect to the transcription system, the present findings add to the other recent observations that reveal the existence of RNAP subunits and transcription factors that are specifically



□□□n □□i r□□l Q,SB□r □tt□□□ ;□ CQ□o□□SF□Jl □□□□□) 5□HK qOG6□□l □□□ □□Jv□□□V□ □vSSJ□□□u2or B□2□,□□J□□□ v□MW□2□r Jt□t g□2□SF□Jl □□□ □ □□HK qO□□ □l □□SuJBvSuM□□ □l □ □,□K □K □□v2□SF□Jl o2□BvSuM□□ □□2l SV r □or □ □,u□9□□□□gv2oSr □SF□Jl □□Jv□□□V dl □□SK M□JJ□□or FSvK t□t Sr □FSv□□ ,□Jl □□2□ku□r □□2□u2□□□FSv□Jv□□□ □Sr 2Jvu□ttSr □□ □□□□□□□□SSJ2JvM□g□ u□2□□□ □□□□FSur□□□ □□uMVJ□K □rJv□ O□□dB6□T6



shared between eukaryotes and Crenarchaeota, along with either Thaumarchaeota or Korarchaeota [20-23].

Authors' contributions

FB,JM and KM performed sequence analysis. FB, KM, and EK wrote the initial draft of the manuscript. JM, BS, and JO wrote the final manuscript. BS, EK, and JO coordinated the study. All authors read and approved the final version of the manuscript.

Acknowledgements

This work was supported by NWO (ALW-Vici project 865.05.001 to JO. JM is supported by the DFG program "GRK 1431/1, Transcription, Chromatin Structure and DNA Repair in Development and Differentiation". KM and EK are supported by intramural funds of the DHHS (NIH, National Library of Medicine). FB and JO would like to thank Finn Werner for helpful comments on the manuscript.

Methods

Sequence analysis

Refseq database at the NCBI [10] was used for PSI-BLAST searches. Database searches were performed using PSI-BLAST [9] with default. We also used the remote homology identification servers for CDD-search [11] and HH search [12]. Multiple alignments of protein sequences were constructed by using MUSCLE program [24], followed by a minimal manual correction on the basis of local alignments obtained using PSI-BLAST [9]. Protein secondary structure was predicted using the PSIPRED program [13].

Maximum likelihood (ML) phylogenetic trees were constructed from the alignment of archaeal RPC34 orthologs (the position used for

Figure 3: Genome context analysis of archaeal RPC34 orthologs. Bold lines indicate possible co-transcription (intergenic region < 100 bp). The genome of *Sulfolobus solfataricus* P2 contains a Gar1 homolog (LocusTaq Sso6830), which is not annotated as a gene, upstream of Sso0944.

a[illegible] t ofar [illegible] f i t [illegible] a [illegible] o [illegible] i l t [illegible] t [illegible] f [illegible] [illegible] r a [illegible] 4-by [illegible] u [illegible] o t [illegible] f [illegible] [illegible] [illegible] [illegible] f ai [illegible] a [illegible] n [illegible] PTOC [illegible] l f [illegible] f [illegible] [illegible] [illegible] [illegible] [illegible] or [illegible] of fr f i t [illegible] h [illegible] fa g [illegible] i [illegible] f [illegible] a [illegible] an [illegible] i [illegible] e [illegible] a [illegible] [illegible] aa [illegible] t [illegible] n [illegible] t f [illegible] [illegible] t [illegible] a [illegible] t [illegible] e [illegible] f [illegible] [illegible] fa [illegible] PTxOR [illegible] [illegible] [illegible] [illegible] f ai [illegible] a [illegible] n [illegible] l [illegible] o [illegible] [illegible] eoi [illegible] r o [illegible] [illegible] fi [illegible] [illegible] i n f r f [illegible] [illegible] [illegible] [illegible] i i fofa [illegible] f [illegible] h [illegible] er [illegible] oR

[illegible] g [illegible] V [illegible] v27 [illegible] SK K [illegible] r J2

[illegible] tt [illegible] fi [illegible] l6 [illegible] ET I [illegible] F [illegible]

[illegible] / [illegible] s [illegible] [illegible] d [illegible] su [illegible] / d [illegible] r s / [illegible] u [illegible] t or d [illegible]

[illegible] SK M [illegible] J [illegible] [illegible] or 2 [illegible] B l e u, [illegible] tt [illegible], [illegible] S F [illegible] [illegible], SK [illegible] l [illegible] J [illegible] [illegible] 6 [illegible] MS M [illegible] 2 [illegible] 2 [illegible] [illegible] MS g [illegible] [illegible] 2 [illegible] [illegible] SK M [illegible], , or B [illegible] B [illegible] r SK d [illegible] g [illegible] r [illegible] [illegible] FS v [illegible] [illegible] g [illegible] vO [illegible] or J [illegible] v [illegible] 2ttr B [illegible] [illegible] gS, uttS r [illegible] O [illegible] l CMS Jl [illegible] 2 [illegible] 2 [illegible] 2 l [illegible] [illegible] [illegible] or B [illegible] r [illegible] V [illegible], dBl J [illegible] S r [illegible] Jl [illegible] [illegible] S v [illegible] B or [illegible] S F [illegible] Jl [illegible] [illegible] u1 [illegible] O S tt [illegible] [illegible] Jv [illegible] 2 [illegible] vdMttS r [illegible] K [illegible] l [illegible] or [illegible] v [illegible] 6 [illegible] [illegible] 2 [illegible] [illegible] [illegible] S r [illegible] [illegible] [illegible] J [illegible], [illegible] [illegible] [illegible] SK M [illegible] [illegible] t g [illegible] [illegible] [illegible] [illegible] , CO [illegible] 2 [illegible] S F [illegible] 2 [illegible] k u [illegible] r [illegible] [illegible] 2 j [illegible] [illegible] SK [illegible] or [illegible] S v B [illegible] [illegible] dtt [illegible] t S r [illegible] [illegible] [illegible] B [illegible] r SK [illegible] [illegible] S r J [illegible] dJ [illegible] Jl [illegible] [illegible] [illegible] Jl Sv2 [illegible] MV [illegible] [illegible] [illegible] dO [illegible] J [illegible] [illegible] [illegible] [illegible] [illegible] vS, [illegible] [illegible] S F [illegible] [illegible] [illegible] ur [illegible] l [illegible] [illegible] [illegible] J [illegible] vd [illegible] [illegible] [illegible] [illegible] [illegible]

On the other hand, one cannot exclude the possibility that the archaeal RPC34 and related TBP are not involved in transcription per se but play another fundamental role in archaeal RNA metabolism. Interestingly, the archaeal RPC34 homologue is present in crenarchaea, thaumarchaea and korarchaea but not in Euryarchaea, the latter containing a more distantly related homologue. This indicates that the RPC34 was present in the last common archaeal ancestor and was later on lost in euryarchaea. The authors suggest from their data that the specialization of RNA polymerase between those transcribing coding and non-coding genes might have evolved already in Archaea and was inherited by Eukaryotes from the "archaeal parents". I previously criticized this notion of archaeal parents, noticing that we don't descend from Apes. Eugene Koonin correctly pointed out that I was wrong since we are apes indeed! However, are Eukaryotes Archaea?? We don't know. It might be that Archaea are reduced proto-eukaryotes? In my opinion, it's still a prejudice to consider that Eukarya evolved from Archaea. I would say that the data presented in this nice paper indicate that the RPC34 protein was present in the last common ancestor of Archaea and Eukarya. As an alternative to the hypothesis proposed by the authors, it could be that this ancestor contained the ancestor of RNA polymerase III and that this protein (but not RPC34) was lost in Archaea (streamlining).

Authors' response

We appreciate the constructive remarks and would like to briefly comment on only two aspects. First, it is hard to agree that "one cannot exclude the possibility that the archaeal RPC34 and related TBP are not involved in transcription per se but play another fundamental role in archaeal RNA metabolism". All we know about these proteins points to direct involvement in transcription, so this seems to be a safe bet. Of course, our suggestion that they are involved specifically in structural RNA synthesis is far more speculative. Second, about the "archaeal parent" of eukaryotes, very briefly, because this issue is far beyond the scope of the paper. Although much of it is semantics, meaningful distinctions can be made. Humans are indeed apes, the third species of chimpanzee by any legitimate criterion used in evolutionary biology. By contrast, eukaryotes are not archaea for the crucial reason that their genetic makeup is an archaeo-bacterial chimera. This is the reason why we find it preferable to speak of the archaeal "parent" of eukaryotes rather than the archaeal ancestor. This logic is not much affected by the exact nature of the archaeal parent – whether it was a typical archaeon or a derived one with some evolved eukaryotic features.

References

1. Landick R: **Functional divergence in the growing family of RNA polymerases.** *Structure* 2009, **17:**323-325.
2. Werner F: **Structure and function of archaeal RNA polymerases.** *Mol Microbiol* 2007, **65:**1395-1404.

) 6 [illegible]

5 6 [illegible]

G 6 [illegible]

R 6 [illegible]

w 6 [illegible]

: 6 [illegible]

P 6 [illegible]

T z 6 [illegible]

T T 6 [illegible]

T (6 [illegible]

T) 6 [illegible]

T 5 6 [illegible]

T G 6 [illegible]

T R 6 [illegible]

T w 6 [illegible]

T : 6 [illegible]

T P 6 [illegible]

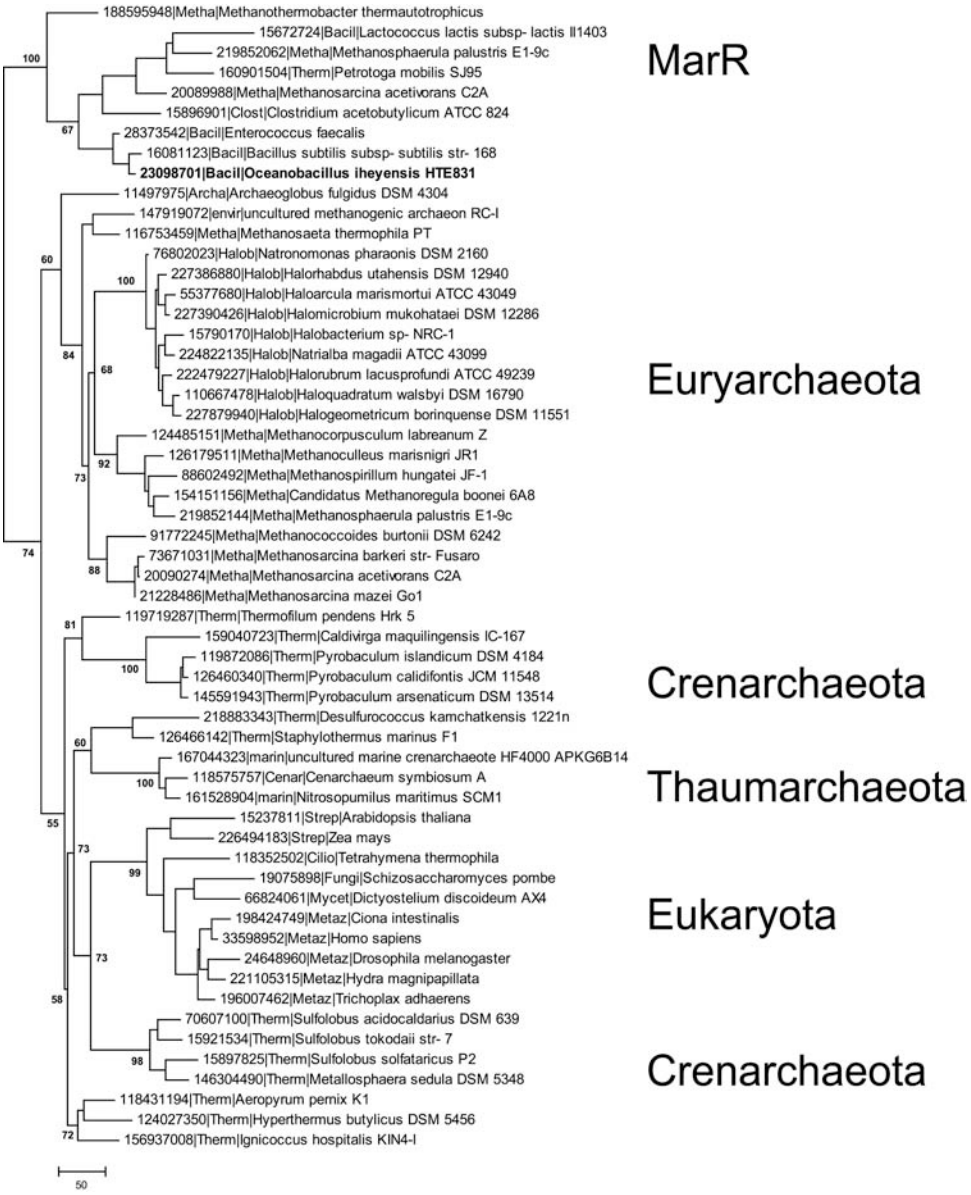
(z 6 [illegible]

(T 6 [illegible]

((6 [illegible]

23. Korkhin Y, Unligil UM, Littlefield O, Nelson PJ, Stuart DI, Sigler PB, Bell SD, Abrescia NG: **Evolution of Complex RNA Polymerases: The Complete Archaeal RNA Polymerase Structure.** *PLoS Biol* 2009, **7:**e102.
24. Edgar RC: **MUSCLE: a multiple sequence alignment method with reduced time and space complexity.** *BMC Bioinformatics* 2004, **5:**113.
25. Adachi J, Hasegawa M: **Computer Science Monographs No. 27. MOLPHY: Programs for molecular phylogenetics.** *Institute of Statistical Mathematics, Tokyo* 1992.
26. Felsenstein J: **Inferring phylogenies from protein sequences by parsimony, distance, and likelihood methods.** *Methods Enzymol* 1996, **266:**418-427.

[illegible]


188595948|Metha|Methanothermobacter thermautotrophicus
15672724|Bacil|Lactococcus lactis subsp- lactis Il1403
219852062|Metha|Methanosphaerula palustris E1-9c
160901504|Therm|Petrotoga mobilis SJ95
20089988|Metha|Methanosarcina acetivorans C2A
15896901|Clost|Clostridium acetobutylicum ATCC 824
28373542|Bacil|Enterococcus faecalis
16081123|Bacil|Bacillus subtilis subsp- subtilis str- 168
23098701|Bacil|Oceanobacillus iheyensis HTE831
11497975|Archa|Archaeoglobus fulgidus DSM 4304
147919072|envir|uncultured methanogenic archaeon RC-I
116753459|Metha|Methanosaeta thermophila PT
76802023|Halob|Natronomonas pharaonis DSM 2160
227386880|Halob|Halorhabdus utahensis DSM 12940
55377680|Halob|Haloarcula marismortui ATCC 43049
227390426|Halob|Halomicrobium mukohataei DSM 12286
15790170|Halob|Halobacterium sp- NRC-1
224822135|Halob|Natrialba magadii ATCC 43099
222479227|Halob|Halorubrum lacusprofundi ATCC 49239
110667478|Halob|Haloquadratum walsbyi DSM 16790
227879940|Halob|Halogeometricum borinquense DSM 11551
124485151|Metha|Methanocorpusculum labreanum Z
126179511|Metha|Methanoculleus marisnigri JR1
88602492|Metha|Methanospirillum hungatei JF-1
154151156|Metha|Candidatus Methanoregula boonei 6A8
219852144|Metha|Methanosphaerula palustris E1-9c
91772245|Metha|Methanococcoides burtonii DSM 6242
73671031|Metha|Methanosarcina barkeri str- Fusaro
20090274|Metha|Methanosarcina acetivorans C2A
21228486|Metha|Methanosarcina mazei Go1
119719287|Therm|Thermofilum pendens Hrk 5
159040723|Therm|Caldivirga maquilingensis IC-167
119872086|Therm|Pyrobaculum islandicum DSM 4184
126460340|Therm|Pyrobaculum calidifontis JCM 11548
145591943|Therm|Pyrobaculum arsenaticum DSM 13514
218883343|Therm|Desulfurococcus kamchatkensis 1221n
126466142|Therm|Staphylothermus marinus F1
167044323|marin|uncultured marine crenarchaeote HF4000 APKG6B14
118575757|Cenar|Cenarchaeum symbiosum A
161528904|marin|Nitrosopumilus maritimus SCM1
15237811|Strep|Arabidopsis thaliana
226494183|Strep|Zea mays
118352502|Cilio|Tetrahymena thermophila
19075898|Fungi|Schizosaccharomyces pombe
66824061|Mycet|Dictyostelium discoideum AX4
198424749|Metaz|Ciona intestinalis
33598952|Metaz|Homo sapiens
24648960|Metaz|Drosophila melanogaster
221105315|Metaz|Hydra magnipapillata
196007462|Metaz|Trichoplax adhaerens
70607100|Therm|Sulfolobus acidocaldarius DSM 639
15921534|Therm|Sulfolobus tokodaii str- 7
15897825|Therm|Sulfolobus solfataricus P2
146304490|Therm|Metallosphaera sedula DSM 5348
118431194|Therm|Aeropyrum pernix K1
124027350|Therm|Hyperthermus butylicus DSM 5456
156937008|Therm|Ignicoccus hospitalis KIN4-I
50
MarR
Euryarchaeota
Crenarchaeota
Thaumarchaeota
Eukaryota
Crenarchaeota

[illegible]

TT:

CHAPTER 7

Summary and general discussion

sy y 4

factors and in Bacteria and Eukaryotes a major group of ribosome assembly factors are GTPases of the translation-factor related class (TRAFAC). There are only a few TRAFAC GTPases in Archaea that serve as candidates for a role in ribosome assembly, as outlined in **Chapter 1**. In **Chapter 2**, one of those archaeal TRAFAC GTPases, the HflX ortholog SsGBP from *S. solfataricus* was studied in structural detail. The crystal structure revealed a two-domain arrangement including a prototypic HflX-domain that probably functions as a nucleic acid binding domain. The presence of the HflX-domain influences the GTP-hydrolysing properties of the C-terminal G-domain. Overall, the enzyme shows biochemical properties similar to other translational GTPases with slow intrinsic GTPase activity and relatively low affinity for GTP.

Chapter 3 describes the interaction of SsGBP with the large ribosomal subunit. Binding assays show that SsGBP similar to its bacterial orthologs binds to the 50S ribosomal subunit both in GDP- and GTP-bound form. Apo-SsGBP shows less stable binding. These results were somewhat unexpected as no conformational changes were observed between the GDP-bound and apo-SsGBP crystal structures whereas GTP-binding was predicted to drive major conformational rearrangements **(Chapter 2)**. In line with the prediction derived from the SsGBP crystal structures, the HflX domain provides a major surface for the ribosome interaction of SsGBP.

In *S. solfataricus* the gene coding for SsGBP is co-transcribed with a gene coding for an archaeal ortholog of the eukaryotic transcription co-activator MBF1. Archaeal and eukaryotic MBF1 orthologs share a conserved helix-turn-helix domain. **Chapter 4** describes the results of a genetic approach for the characterization of the *mbf1* gene. It is shown that, at least under the conditions tested, *mbf1* is not essential in the crenarchaeon *Sulfolobus solfataricus* and gene deletion caused only a mild phenotype and little change on transcriptome level. In **Chapter 5** it is shown that the archaeal MBF1 shows properties that differ from the published data for its eukaryotic ortholog. While no evidence was found for an interaction with the transcription machinery, the archaeal MBF1 ortholog does bind to ribosomes engaged in translation via the 30S ribosomal subunit. Intriguingly, the helix-turn-helix domain of archaeal MBF1 provides the major binding surface for the interaction with the 30S ribosomal subunit.

2 unsnu02

gyeymydggdrseaealmdrgnlaarmgdrm
edhdhdldhaldlSdoregheyourayeadurerageyyesaeed
rmagrgeuladerafdlSdoregrdraeedrmBr
reoiaA dzuydaeH0h DBragnlaeuohegugrlaudds eddeSarsa dzuyouwldorehdadrehd
lmdrehrardrehpdydgaemrarayouedrlfiddudaeaeReragrdzuderurgerlSdoreydreagairuaa
eydrggaarraaeaardmeraldedderaeggldggrafcmyugdeeeelmdzuerryrgrmdgrderur(GDeedaduedagdyeraurareraererdzu)derur("DrdederurgdzG"derurgzG"lvDaayrumdgglaeG"derurgzG(zDaalm–adadrefaeRerarreggdzuderurguydsaudzydelrreylee AdzuydaeHrlSdoregfarerdaeddgeaRaaggeeelmdrehlgeglSdoregrrareraeaeygueaeaeraraudeaeraraudeaeraeraeygueaeraeraeygueglaeasdgeggueglaaauredmdzuf
i aueoreRdgeelmdzuaSrdzuarmg(Oh- DugrursgrdeddauogrreelddarrdudzdaeRdzuderurgrlSdoreydreagf aeegrraecehaaeddgeaRaaggeee) gyydaeuomggaaalmdrehlmdrehlgeglSdoregrrareraeaeygueaeaeraraudeaeraraudeaeraeraeygueaeraeraeygueglaeasdgeggueglaaauredmdzuf
ir egeggeggydri gi aardlaeraudlda esaddzuaelSdoreeee

subunit [13]. Thereby one has to rely on the only structure for the small ribosomal subunit available, this is the 30S ribosomal subunit from the bacterium *Thermus thermophilus*. Recent advances in cryo-EM studies on eukaryotic ribosomes from *Thermomyces lanuginosus* allowed to include all ribosomal proteins with bacterial orthologs into the eukaryotic ribosome model, but still those specific for Eukaryotes and Archaea are missing [14]. No such model exists for the 70S archaeal ribosome or the 30S ribosomal subunit. Any co-crystallization experiments for the MBF1:30S ribosomal subunit-complex would therefore be confronted with the same problem. It might thus be worth to test whether MBF1 is also able to bind to bacterial 30S ribosomal subunits. Binding to the bacterial 30S ribosomal subunits could occur when the structural determinants for MBF1 binding on ribosome are conserved between Archaea and Bacteria. In order to test whether this is the case, other biochemical data characterizing the interaction of MBF1 with its cognate 30S ribosomal subunit have to be generated first. In addition, structural information for the archaeal MBF1 protein itself is required. To this end the ongoing NMR studies and chemical probing experiments are likely to set the basis for future studies allowing to model the MBF1-30S ribosomal subunit interaction.

The discovery of the archaeal RPC34 ortholog described in **Chapter 6** follows several recent discoveries concerning orthologies between components of the different eukaryotic RNA polymerase systems [15] as well as orthologies between the archaeal and eukaryotic transcription systems [4, 16]. TFIIEα, the eukaryotic ortholog of the archaeal transcription factor TFE, has recently been proposed to be a paralog of the RNA polymerase III subunit RPC82 that forms a heterodimer with RPC34 [17]. Although the evidence from bioinformatic analysis is rather weak, the authors also proposed that RPC34 is a paralog to TFIIEβ. This further suggests that a gene duplication of a heterodimeric transcription factor in the ancestor of today's eukaryotes led to TFIIE in the RNA polymerase II system and the RPC34/RPC82 sub-complex of RNA polymerase III. The RPC34/RPC82 sub-complex of RNA polymerase III could thus be the result of a "transcription factor capture" event where a loose transcription factor interacting with the RNA polymerase became incorporated as a stably associated RNA polymerase sub-complex. The fact that Archaea possess orthologs to TFIIEα and RPC34 [7, 18, 19] raises the question at which time point during evolution the heterodimerisation did occur. A key question is whether there is an interaction between TFE and the

G" rderur a d hra grhs eedtegeu rdedagae aedeeraf egedoaedl a eeeeeeeeeeugeeegnl a uzei uzeddgegedra ddrragdi eera es a eeddu G" rderur a l Sdoree G" s urreeeeeeeeee

been proposed to function in ribosome assembly as well [31]. Thus, the eukaryotic HflX orthologs might have similarly functions in protein synthesis in chloroplasts and mitochondria.

Much progress has been made regarding the interaction of many TRAFAC GTPases with ribosomes or free ribosomal subunits, sometimes to a level where defined precursors of ribosomal subunits were identified as binding target. Thereby the precise time point for the functioning of these GTPases during ribosome assembly could be identified. Nevertheless, in the vast majority of cases the working mechanism of GTPases is not well understood. Working models include the recruitment of extrinsic factors that are for instance involved in ribosome assembly, to the ribosomal subunits. In such a scenario the nucleotide-bound state of a GTPase will be crucial to regulate binding to such an extrinsic factor and its release on the ribosomal subunit. The ribosome-stimulated GTPase activity of HflX fits well into such a model and could trigger the release of a factor. However, the phylogenetic distribution of HflX GTPases with many different organisms in all three domains of life missing the *hflX* gene makes it unlikely that this GTPase family functions in factor recruitment. No factor with a phylogenetic distribution similar to HflX could be identified in a bioinformatical analysis. This means that either in several instances the recruitment of this extrinsic factor to the ribosome became independent of HflX, or that HflX GTPases act on different extrinsic factors pointing to functional adaptation during evolution. Nevertheless, a screen for protein interactors of HflX GTPases could provide important insights into their physiological function, especially regarding the possibility that the *S. solfataricus* HflX ortholog is involved in the release of an extrinsic factor from the 50S ribosomal subunit.

For all three proteins under investigation in this thesis – the archaeal RPC34 orthologs in general and MBF1 and SsGBP (HflX) from *S. solfataricus* – no biochemical information was available at the beginning of this project. At about the same time that we determined the structure of SsGBP and discovered its interaction with the 50S ribosomal subunit, the first functional characterization for bacterial HflX representatives became available [21, 22, 32]. Although these data do not suffice to determine the physiological function of HflX GTPases, they do suggest that the function might have been conserved in some extent between archaeal and bacterial HflX GTPases. Apart from the relatively well-characterized translation initiation and elongation factors, SsGBP remains the only archaeal TRAFAC GTPase with a predicted function in

[illegible]

[illegible]

[illegible]

kY [illegible]

vY [illegible]

KY [illegible]

hY [illegible]

LY [illegible]

WY [illegible]

IY [illegible]

EY [illegible]

VY [illegible]

kMY [illegible]

kkY [illegible]

12. Hendrick JL, Wilson PG, Edelman, II, Sandbaken MG, Ursic D, Culbertson MR: **Yeast frameshift suppressor mutations in the genes coding for transcription factor Mbf1p and ribosomal protein S3: evidence for autoregulation of S3 synthesis.** *Genetics* 2001, **157:**1141-1158.
13. Sharma MR, Barat C, Wilson DN, Booth TM, Kawazoe M, Hori-Takemoto C, Shirouzu M, Yokoyama S, Fucini P, Agrawal RK: **Interaction of Era with the 30S ribosomal subunit implications for 30S subunit assembly.** *Mol Cell* 2005, **18:**319-329.
14. Taylor DJ, Devkota B, Huang AD, Topf M, Narayanan E, Sali A, Harvey SC, Frank J: **Comprehensive molecular structure of the eukaryotic ribosome.** *Structure* 2009, **17:**1591-1604.
15. Werner F: **Structural evolution of multisubunit RNA polymerases.** *Trends Microbiol* 2008, **16:**247-250.
16. Koonin EV, Makarova KS, Elkins JG: **Orthologs of the small RPB8 subunit of the eukaryotic RNA polymerases are conserved in hyperthermophilic Crenarchaeota and "Korarchaeota".** *Biol Direct* 2007, **2:**38.
17. Carter R, Drouin G: **The increase in the number of subunits in eukaryotic RNA polymerase III relative to RNA polymerase II is due to the permanent recruitment of general transcription factors.** *Mol Biol Evol* 2009.
18. Hanzelka BL, Darcy TJ, Reeve JN: **TFE, an archaeal transcription factor in Methanobacterium thermoautotrophicum related to eucaryal transcription factor TFIIEalpha.** *J Bacteriol* 2001, **183:**1813-1818.
19. Bell SD, Brinkman AB, van der Oost J, Jackson SP: **The archaeal TFIIEalpha homologue facilitates transcription initiation by enhancing TATA-box recognition.** *EMBO Rep* 2001, **2:**133-138.
20. Comartin DJ, Brown ED: **Non-ribosomal factors in ribosome subunit assembly are emerging targets for new antibacterial drugs.** *Curr Opin Pharmacol* 2006, **6:**453-458.
21. Polkinghorne A, Ziegler U, Gonzalez-Hernandez Y, Pospischil A, Timms P, Vaughan L: **Chlamydophila pneumoniae HflX belongs to an uncharacterized family of conserved GTPases and associates with the Escherichia coli 50S large ribosomal subunit.** *Microbiology* 2008, **154:**3537-3546.
22. Dutta D, Bandyopadhyay K, Datta AB, Sardesai AA, Parrack P: **Properties of HflX, an enigmatic protein from Escherichia coli.** *J Bacteriol* 2009, **191:**2307-2314.
23. Engels S, Ludwig C, Schweitzer JE, Mack C, Bott M, Schaffer S: **The transcriptional activator ClgR controls transcription of genes involved in proteolysis and DNA repair in Corynebacterium glutamicum.** *Mol Microbiol* 2005, **57:**576-591.
24. Morimoto T, Loh PC, Hirai T, Asai K, Kobayashi K, Moriya S, Ogasawara N: **Six GTP-binding proteins of the Era/Obg family are essential for cell growth in Bacillus subtilis.** *Microbiology* 2002, **148:**3539-3552.
25. Qin Y, Polacek N, Vesper O, Staub E, Einfeldt E, Wilson DN, Nierhaus KH: **The highly conserved LepA is a ribosomal elongation factor that back-translocates the ribosome.** *Cell* 2006, **127:**721-733.
26. Britton RA: **Role of GTPases in bacterial ribosome assembly.** *Annu Rev Microbiol* 2009, **63:**155-176.
27. Leipe DD, Wolf YI, Koonin EV, Aravind L: **Classification and evolution of P-loop GTPases and related ATPases.** *J Mol Biol* 2002, **317:**41-72.
28. Barrientos A, Korr D, Barwell KJ, Sjulsen C, Gajewski CD, Manfredi G, Ackerman S, Tzagoloff A: **MTG1 codes for a conserved protein required for mitochondrial translation.** *Mol Biol Cell* 2003, **14:**2292-2302.

vVY 2 9 f l 029 f y o a0 so a f uy f y

f l . u s RA Tu E 7u E EA E 7u 7 A S Sup u 7 mT S

SS Tm A E 7 7g P 7gE T SCSo vMV/fl : cKl VtKVMY

KMY S u9eu f uGG r f 0r f 0l Sa2 Sl . 7 C ERSCS p H S t u f S

E A CT Æ 7E RA SuS T S T mE T CT P 7gE y ÆRP Tuo

vMV/fbdkKI KtkKEKY

KkY u2 f S 4e f n CuS 7 7 S Tur S u R7Eu CT ES CP TuS

n Cu u D5 7C ESEP A Sr r TCuo vMVhf BNdhEkthEVY

KvY 2 f Suy 0r f S f y e f hunS f si f e e a f ea S .doEAC

Cu 7 uSh Cu D5 7C ESEP A Sr r TCuSCT R7S T E Tr A Em S o o *o* vMV/fl Ilticv/MktivMLY

SUPPLEMENTS

?e ?nk5 n1

?D??lrD??an DrvDv?ë?v?dDB?mbëDv?R???BDlrsvf??l?d?D??Dv??co?lEaBDv?R?ocvvDv?ë?v?Dmo??l?av?Dl–?dDr?Dv?, al?Dv?ah???–r–?ë?v??D?, riuD?, ??lah??D?DvDB?–?dD?rvealn ?B?D? ?, al?B?ëDl, DloB?B?i?Dv–?dDB?oahrslDv?'lDhm??B?Déf??X?buDv?'B?v–?lrhB?DéfDv??ëDlB?mbv?B?v–?aB?Déë?v?????.?Dv?dDB??mDn DDv?rioB?dDB?Dlah?????B???co?lEaBDv?ëaal??DuD?hla?D—Dv?n ??drvD–?D?lcroDv??rD?dan aaa ?univn DB??D?ëDl Dmio??lD??l?d?D?mb?n ??drvD–.??rnDl?aal?avB–Bav?dDB?r?DD???B???v?dDB??D rv?ë?v??D?Dco?lEaB–?dD?ëamB?D?DDv?D?l?d?D?mh?lDvBM?B??B

?mB??aDoiD? ??B?aëDl?ëDl–?dmmDv?D??–hD?BDv?ë?v?B?v–?aB?D?rv??l?d?D??rv?dDB??mDn DDv?Dv?dDB?al ?vr–n D??????????? ?????????????v?dDB??riuav?Dl.??Dv?laaB???vB?mhDr; mBBDv?r–??rlD?B??DBlaooDv?ri?dDB?B?v–?aB?Dhla?D–??m?e??BalDv?rD?dDB?hla?D–?ë?v??D rv?BaB?Drv??D cmDlDv?a?m?lr?a–an rhD?Dr; mBBDv?rD??DDm?cmBn ?oDv?ë?v?dDB??r?a–aan f??D?n ??drvD??rD??D?Dr; mB-EvBdD–D?o?B?mñ–DDlB??ëDldDB??mDn DDv?univ??DuD?Dr; mBBDv?dDDm?-BDlo?D?av–DlëDDl??Dv?DDv?laaB??DDm?ëv?dDv?r–?BDlc ?BD?ërvDv?rv??mmD?ëaln Dv?ë?v?dDB?mbëDv.?? DDl??v?Dl?D?r; mBBDvuniv??DBlaooDv??ri?ëDl–?dmmDv?D??ri?DdalDv?D?hla?D—Dv.??Dv?ë?v??DuD?hla?D—Dv?r–?dDB?ëaln Dv?ë?v?dDB?lr?a–aan ??aal?dDB??—Dn ?mDlDv?ë?vlr?a–an ??m?????Dv?lr?a–an rhD?Dr; mBBDvf??D?ua DvaDn ?D?lr?a–aan ??ra DvD–D.

?D?lr?a–aan ??ra DvD–D??D rvB?h DB??D?–EvBdD–D?ë?v?lr?a–an ??m?????'l????é??aal?DDv? ?????hamÑn Dl?–D.??co?lEaBDv?dD??Dv?B; DD?????hamÑn Dl?–Dv?'?Dv????éf??rD?–hD?r??m?eaal??D?B?v–?lrhB?D?ë?v?lr?a–an ??m?????Dv?v?Dl D?–B?c?BclDmb?????–?; al?Dv?D?lcroB??mB??DBDoDvB??B??DuD???"–?vrD?ëDl?Dl?, al?Dv?ëDlB??m?v?l?Dr; mBBDvn ??l??rl D?B?, al?Dv? D?lcroB??m?ri?ëaal?DDm?ac, –BDDv?ë?v?dDB?lr?a–aan .??v?BD Dv–BDmv ?BaB??co?lEaBDv?dD??Dv??ll?d?D??Dv????BDlrsv?n ??l?yyv?–aalB??hamÑn Dl?–D.

?D–?dlrieB??D?r?DvB?X??B?D?ë?v?DDv??l?d?D??mDr; mB??B?DDv?alBdamaa ?r–?ë?v??D?Dco?lEaB–?dD?hamÑn Dl?–D?-cDDvdDr?4-x.??DuD?–cDDvdDr??, Dl?DDl?Dl??D–?dac, ??m–?cvrDo?ëaal?dDB?Dco?lEaB–?dD?hamÑn Dl?–D?DB?dDeBDDv??Dm?v lrioD?lam?ri?dDB??r Dl?Dv?ë?v?dDB?hamÑn Dl?–D?v?l??D?mDn DvD?B?v–?lrhB?De?BalDv?rD??v?hlan aBDl?D?rD?Dv? D?av?Dv?univ.??c–??D?ëDl, ??dBv ?r–??B??D?4-xalBdamaa ?rv?l?d?D?aao??ri?D?B?v–?lrhB?D? ?DBlaooDv?r–.??Dv?v?m–D?ë?v??D?Dvaan ?avBDAB?ë?v?dDB?l?d?D?mb?4-x Dv?rv?ëDl–?dmmDv?D?l?d?D?mb?Dvan Dv?n ??oB?

?cr?Dmio??B??mB?Dv?ë??o?v??D??cclBmn B?ë?v?DvDv?rD??DBlaooDv?univ?ri??D?h a?rX??B?D?Dv?dDB?ovrhhDv?ë?v?–?lc?BclDmb????"–.??D?–hD?r?m–?B?D?ë?v?

RNA polymerasen voor eiwitcoderende en structurele RNA's heeft dus misschien zijn oorsprong in de "archaeal parent" van de Eukaryoten met een sleutelrol voor RPC34.

Talrijke factoren controleren de ribosoom biogenese. Een van de grootste groepen van zulke factoren in Bacteriën en Eukaroten bestaat uit GTPasen die tot de translatiefactor klasse (TRAFAC) behoren. Er is maar een klein aantal van TRAFAC GTPasen in Archaea te vinden dat in aanmerking komt voor een functie bij het in elkaar zetten van het ribosoom, hetgeen wordt uitgelegd in **hoofdstuk 1**. In **hoofdstuk 2** wordt de structuur van SsGBP, een GTPase uit *S. solfataricus* behorend tot de HflX familie van TRAFAC GTPasen, beschreven. De kristalstructuur toont aan dat het eiwit bestaat uit twee domeinen waaronder een domein dat wij het HflX-domein hebben genoemd omdat dit domein uniek is voor deze soort eiwitten. Het HflX-domein kan waarschijnlijk nucleïnezuren binden en het lijkt de katalytische eigenschappen van het C-terminale GTPase-domein te beïnvloeden. Over het algemeen zijn de biochemisch eigenschappen van dit eiwit vergelijkbaar met die van andere GTPases die bij de translatie betrokken zijn zoals een lage GTPase activiteit en een lage affiniteit voor GTP.

Hoofdstuk 3 beschrijft de interactie van SsGBP met de grote ribosomale subeenheid. Zoals eerder voor bacteriële HflX GTPasen al is gevonden, bindt SsGBP aan de grote ribosomale subeenheid zowel met GDP als ook met GTP gebonden. De nucleotide-vrije vorm van SsGBP is ook in staat aan de grote ribosomale subeenheid te binden, maar op een minder stabiele manier. Dit was een verrassend resultaat omdat eerder in de kristalstructuren geen veranderingen in de conformatie van het eiwit gevonden werden tussen de vorm van SsGBP die GDP gebonden heeft en de nucleotide-vrije vorm, maar het voorspeld wordt dat het binden van GTP grote structurele wissels in SsGBP zal veroorzaken (**hoofdstuk 2**). Zoals voorspeld op basis van de kristalstructuur van SsGBP levert het HflX-domein de grootste bijdrage voor de interactie met de grote ribosomale subeenheid.

In *S. solfataricus* wordt het gen dat codeert voor SsGBP samen met een andere gen afgelezen. Dit tweede gen codeert voor een eiwit dat gerelateerd is aan de eukaryotische transcriptiefactor MBF1. De archaeale en de eukaryotische MBF1 orthologen delen hun zogenoemde helix-turn-helix-domein. **Hoofdstuk 4** omschrijft de consequenties van een disruptie van het *mbf1* gen. Het gen lijkt niet essentieel te zijn in *S. solfataricus* en de disruptie van het gen heeft slechts een mild fenotype als gevolg. Het was verrassend dat

aao dDB Bv–lmBaan f D–n Dv–BDmv ëvdDB BaB hn v D Dmfh l
, Drvr tDlv Dlrv avDl B

v , al Bv Dvaav B D Dr Dv–dhhDvëvdDB
l d D mh C cr Dmioë Dl–dmmDvëv D D–dl Dë Dv Dr Dv–dhhDvëv
dDB Dco lEaB–dD C. Dv vBDl B Dh DB D Bv–lmB Dn drv D, Dl v nDB
Dëav Dvfh l dDB Dr, rB rv B v D DomDrv D lr a–an mh –c DDvdDr Bi Dv–
B v–m B D. DB r–ahn DloDmio B dDB dDmAlBclvIdDmAl an Drvëvd DB
l d D mh C f B aao v D Dco lEaB–dD C Bv Dër v Dv r–f dDB laaB–B D
v DDmfl DDeBf v D v BDl B Dh DB dDB lr a–aan .

?

Zusammenfassung

Die drei Domänen zellulären Lebens auf der Erde – Bakterien, Archaeen und Eukaryonten – lassen sich aufgrund der Art und Weise unterscheiden, wie die genetische Information, die in der DNA enthalten ist, während der Replikation, Transkription und Translation verarbeitet wird. Im Allgemeinen scheint es so zu sein, dass Eukaryonten für diese Prozesse Maschinerien verwenden, die homolog zu den entspechenden archaealen Gegenstücken sind. Dadurch entstand die Idee, dass am Beginn der Evolution der Eukaryonten ein "archaeal parent" stand, vom dem diese Maschinerien abstammen.

Die vorliegende Arbeit beschäftigt sich mit verschiedenen Aspekten der Translation in den Archaeen im Allgemeinen und in dem Organismus *Sulfolobus solfataricus* im Besonderen. Viele Proteine spielen eine direkte Rolle bei der Translation, z.B. als Faktoren, die von der Initiation bis zur Termination den Translationsprozess kontrollieren, oder als integraler Bestandteil der Ribosomen, welche die Eiweißsynthese katalysieren. Alle diese Proteine sind generell sehr stark konserviert und gehören in ihrer Mehrheit zum elementaren Bestandteil allen zellulären Lebens. Andere Proteine widerum, spielen eine Rolle bei diversen Prozessen, die indirekt an der Translation beteiligt sind. Dazu gehört schon das Zusammensetzen von ribosomaler RNA und den ribosomalen Proteinen zu intakten Ribosomen, die sogenannte Ribosomenbiogenese.

Dieser Prozess des Zusammensetzens beginnt schon während der Transkription der ribosomalen RNA durch die RNA Polymerase. Eukaryonten besitzen zwei RNA Polymerasen (I und III), die speziell der Transkription von ribosomaler RNA und anderer struktureller RNA (die nicht in Protein übersetzt wird) dienen. Im Gegensatz zu den Eukaryonten besitzen Archaeen und Bakterien nur einen einzigen Typ RNA Polymerase.

Kapitel 6 beschreibt die Entdeckung eine archaealen Orthologs der eukaryontischen RNA Polymerase III Untereinheit RPC34. Dieses Protein galt vorher als spezisch für die RNA Polymerase III. Es nimmt eine wichtige Funktion bei der Rekrutierung der RNA Polymerase III durch die basalen Transkriptionsfaktoren ein, die an die Promotorregion eines Genes gebunden sind. Aufgrund dessen scheint es logisch, dass das archaeale RPC34 Ortholog ebenfalls in der Transkriptionsinitiation eine Rolle spielt. Die Analyse des Genomkontexts des Gens, das für das archaeale RPC34 Ortholog kodiert, in verschiedenen archaealen Genomen ergab, dass das Gen häufig Cluster formt

n nB Dv Dvf nD Dr Dl lauD--nDlcv cv a rXurDlcv -BcoBclDmDl –
DrvD amD–hrDmDv. nD ëamBav Dl ëDl–dnD DvDvDcoIEavB–dDv
amfin Dl–Dvf–hDurfm-nDlB ce nD EvBlD–D –BcoBclDmDl a Dl laBDrvI
oa nDlDv Dl f–fabj vvBD Dn v dl dl Dvl–hlcv av Dn tDl l dDD nfh lDvBWV
Dl co IEavBDv d Dvf, a Dr Dn 4-x lBdam DrvD, r dB- D amD
ucokn D.

nDmD oBalDvbavBamnDlDv Dv lauD--Dl r a–an DvIra DvD-D.fv
oBDlnDvctv co IEavBDv+BDmDv –Dvc–Dl lv-nfBav-eboBall nfa--D
' é nD ljzBD lchhD nD-Dl oBalDv. v l d lDDv Xv Dv–r d vcl
, Dvr D –Dvf nD ma v r BDv egl rnD r a–an Dv ra DvD-D
v D-DdDv, Dl Dvbj vvDv. rvD Dl–r dBdnDlg Dl rnDBDB . D
D-dlDr B nD B coBcl ëav –f Drv Dl –Dc– uc
Dl Xfal n mnD rvvDld nal Dl –Dv Ddj lB nD lr–B ma B coBcl
uDr Bf––– –c–u, Dr an kvDvD-BDdB nD lBDln rv mD an kvDdl Dv
, nl Xfal an kvD Dv vvBf nD-D an kvD v-dlDrvDv vcl av Xfal –Dv
uc Xv Dv r–B nD Xfal an kvD nDvB, dl–dDrvm d Dl rv cv ëav Xfal
–Dv v comDrv–kclDv. D–, DrBDlDv+dDrvB nD lk–Dvu Dl Xfal an kvD
nD oB mfBr–dDv r Dv–d eBDv Dl lBDln rv mDv –D an kvD uc
DrvXm--Dv. nD oB mfBr–dDv r Dv–d eBDv ëav –kdvDmv m
mnDn DrvDv DvDv ërDmDl v Dl l –Dv n nB cvoBravDv rv Dl
l v–nfBav–n –dDrvDlnDf, DrfdD–r d DvDlDmfcl d DvD lDrfBë Dlrv D
eXvmBkBglcv DrvDvrD lr D E lamf–Dl BDc–uDr dvDv.

D dlDr B nD vBDloBravu, r–dDv –cv Dl lazDv
lr a–an mDv vBDlDrvdDrB dvm d, nDucël –davegl oBDlDmDm Xfal
–Dv D–dlnD Dv, cl Df rv DB –a, adm m D cv DvDvf, nD
c d m D cv DvDv c–B v v nD lazD lr a–an mDv vBDlDrvdDrB nD
vcomBaB elDrD aln ëav –o vv D Dve ma v nD lazD lr a–an mDv
vBDlDrvdDrB rv Dvf mDl rv –a–B nD rv cv , Dvr Dl +B m f nD-D- l D vr–
, l g Dll –dDv f D Drv Dl –DrB-r rv Dv lr–B ma B coBcl Dv oDrvD
avealn Bav–kv Dlcv Dv u, r–dDv Dl vcomBaB elDrDv cv Dl
D cv DvDv aln ëav –Da dBDB, cl Dvf v Dl Dl –DrB- ce –r– Dl
lr–B ma B coBcl Dv +an dD avealn Bav–kv Dlcv Dv ëaldDl D– B, cl Dv egl
nD D cv DvD aln ëav –é. nDëaldDl D– Bf-hrDmBfnD
Xfal an kvD Drv D, r dB- D amD Dr Dl rv cv v nD lazD lr a–an mDv
vBDlDrvdDrB

v , nl –Dvf –egl –oa nDlBfuc–n n Dvn nB
DrvDn v Dl Dv Dv Bv–olr nDlBf –egl Drv laBDrv oa nDlBf, Dfd D-

verwandt ist mit dem eukaryontischen Transcriptionsfaktor MBF1. Die archaealen und eukaryontischen MBF1 Orthologe besitzen beide C-terminal eine sogenannte Helix-turn-Helix Domäne. **Kapitel 4** beschreibt den Phänotyp, der aus der Gendeletion von *mbf1* resultiert. Das *mbf1* Gen ist nicht essentiel in *S. solfataricus* und wir konnten nur einen relativ mild ausgeprägten Phänotyp beobachten. Überraschenderweise weist auch das Transkriptom, die Gesamtheit der RNA in den Zellen, kaum Veränderungen auf. In **Kapitel 5** wird das MBF1 Protein von *S. solfataricus* biochemisch charakterisiert und es wurde deutlich, dass sich die Eigenschaften von den beschriebenen Eigenschaften eukaryontischer MBF1 Proteine unterscheiden. Es wurden keine Interaktionen mit dem Transkriptionsapparat gefunden, hingegen bindet das archaeale MBF1 an die kleine ribosomale Untereinheit während der Translation. Bemerkenswerterweise findet die Bindung an die kleine ribosomale Untereinheit hauptsächlich über die C-terminale Helix-turn-Helix Domäne statt, welche auch im eukaryontischen MBF1 Protein vorhanden ist.

anvo m 1e nZ0

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

of submitting and resubmitting the manuscript for the crystal structure of SsGBP, I was glad to know that you were always there to deliver in no time the modified figures and manuscript parts. Rie, thanks for all your support in form of chocolate, food, and Panda bear photos on my computer when I was in a bad mood.
Some people that helped significantly to make me having a great time here: Paweł and Nicolas (thanks for the good company ~~all~~ most of the time), Matthijs (you and Stan did your best to corrode my German political correctnes. Thanks for your help with the Dutch summary), Jaapie (I am looking forward to spending some more time at Ashburn, VA!), Mark van Passel (for very valuable comments on manuscripts and applications), Harmen (for betting all the time that "die Mannschaft" gets defeated", they did not).
I wish also to acknowledge my two students, Violeta and Daan. Violeta, you did your very best to become the hardest working student ever. Daan, in spite of all the bad luck we had with the heterologous expression of five different proteins you did not loose your motivation. You both were a great help during the difficult times of the SsGBP and MBF1 projects!
Thanks to all my fellows from the organization comitte of the PhD trip 2009, Matthijs, Marcel, Thomas, and John (R.). Although the organization was very time consuming, I do not regret that I did join the team. It was fun and probably the best experiences during my time as PhD student.
All old and new BacGen members not mentioned yet: Jasper (W.), Pierpaolo (my italian language teacher), Edze, Ans, Sjon, Serve, Suzanne, Marco, Mark L., Mark M., Ronnie, Bram, Magnus, Kathrin, Colin, Marke, and Wouter, thanks for the good atmosphere at BacGen!
All the technical and administrative staff at the Laboratory of Microbiology, Wim, Hans, Ton, Francis, Anja, Renee, and Nees, thanks for making things run so smoothly there. Especially I wish to thank Wim who always had to deal with my anger after another failed ultracentrifuge run.

Ich möchte auch meiner Mutter und Wilhelm danken für all die Unterstützung und das Verständnis während der letzten vier Jahre.
Anja, nach so vielen Wochenenden, die ich dich ins Labor geschleppt habe, und nach all den durchkreuzten Urlaubsplänen, bin ich dir dankbar für all deine Liebe.

□t MM□P □Lpu

BJ R

About the author

Fabian Blombach was born on 7th of August 1977 in Frechen, Germany. After high school graduation at the Norbert Gymnasium Knechtsteden, he performed alternative civilian service for thirteen months. He enrolled for studying environmental sciences at the University Duisburg-Essen (formerly University Essen) in October 1998. After the intermediate diploma in October 2000, he became interested in Microbiology and Molecular Biology and was offered a position as student research assistant by Prof. Bettina Siebers at the laboratory of Microbiology headed by Prof. Reinhard Hensel. The project he worked on aimed to characterise the transcription machinery of the hyperthermophilic archaeon *Thermoproteus tenax* and its regulatory potential. He stayed in the laboratory of Microbiology to conduct his diploma thesis research under the supervision of Prof. Bettina Siebers on the establishment of an in vitro transcription system for *T. tenax*. He graduated in March 2005. After staying for another 7 months as research assistant at the laboratory of Prof. Reinhard Hensel and Prof. Bettina Siebers, he became PhD student in the group of Prof. John van der Oost, Wageningen University, Netherlands where he worked on a project entitled "Back to the future: unvravelling eukaryal-like control networks in Archaea".

9 Z v t c r 8 vn0

c i 6 c n i 6 v e 2 i 6 3 v c n O 6 6 k n 3 O 6 v 3 n d n 6 k n c k m 6 a 6 c 6 9 2 i 6 v e 3 i 6 k n 3 v O Z 6 p t p t E u P u E l L u k 6 tiGdK, P KBJ 8

i 6 a 3 v e i 6 a e v e i 6 k n 3 v O Z 6 c 3 a e Z v e e i a 6 T P E t u S E S T L u E L p T E u p u E v M u E L s u E E p u t E u i L u E L E L g t P p d L E L u k F, , K6 G d J z P X J J z K, 8

v n 9 h 1 6 v e 2 6 c 6 3 v c n O 6 k n 3 v O Z 6 E U E P t U M T p L T E r L T p E T r L u f F, , X6 h t i c P F PK8

v e 2 6 a 3 v k 6 3 3 v 6 9 3 O 6 v v n 9 h 6 k n 3 v O l l l E L E L E T p E U E t S T t i l M E U t P T u t u t L p h i L T L T E p L t P T E p t u u u u M U g l E L E M E U t P T u u E T E L L L E L d E L u t u L T k F, , X6 i d X8

c n 1 6 c 6 v 6 v e 2 6 k n 3 v O Z 6 9 2 6 n 2 n 1 6 v 7 t L l E L U u p t t t E L t t t E T E U h i u t t p L E L M T E p L T E P t U E U t u u E U p E T t u A P P T E p u t t P F, B, 69 h t t 2 0 0 8

9 9 3 O 6 Z d 6 h i 3 Z l 6 S 2 n 6 m 3 O 6 m m 6 v e 2 6 9 Z 9 k v n 6 m Z d 9 h 6 l 3 t t 9 O 6 d v 0 9 n c O 6 m n O 6 e t t t 6 a 3 v k 6 m n a 6 Z t t 2 n 8 2 c O Z 3 n O m 7 E P M U p L E P u N t L E E T P E M T E p t u p L v c M t t u E U U t s T u l U p t t M u p T L E T L T E p k

3 3 v 6 v e 2 6 v n 9 h 1 6 S 2 n 6 9 h a e n n 6 O n 3 6 n O m 6 9 3 O 7 E U E P t U M E M T U E u L T P E M T E p t u p L v k

v e 2 6 n l 6 n 9 2 3 O 6 6 3 Z 2 e 6 v 3 g c 9 v k v 6 c 6 v n 9 h 1 6 4 e 6 9 a e 6 3 9 Z 6 2 3 9 O Z v c v c 6 3 v c n O 6 6 e 6 n 6 v n 9 6 k n 3 v O Z r E T p E U E T E P L u p E p u P U U t t L p t T L p t L u U l E L k

v e 2 6 n l 6 v 3 g c 9 v 6 o 3 Z O 6 3 9 Z 6 n 6 v n 6 2 3 9 O Z v c v c 6 v n 9 6 k n 3 v O Z L q t t M L u p E p b d E u E P U U t t L p t L u U l t U E l e t L u p p u k

n l 6 v e 2 6 Z O c e 9 6 m 3 O 6 9 3 v 9 n 0 a 9 6 v c 1 v k a 9 n 6 4 e 6 Z v e 9 6 v v k 3 O 6 3 v c n O 6 k n 3 v O Z L E p t M L u t u E r T E P p T L u T E M E L k

v e 2 6 3 v c n O 6 k n 3 v O Z u u P P U L p T u U E P d M E p L l P U T E U u E T p T L u U l E L p E T o T U p u u u u k F, BB6 h H

i Z 0 O Z v 3 O v n Z 9 c Z g c h r h

Bj ,

VLAG graduate school activities

Discipline specific activities

- 3 months stay at the laboratory of Prof. Paola Londei, Universita La Sapienza (Rome, Italy) in 2008 and 2009#
- 4 days stay at the Institute of Structural Molecular Biology, University College London (London, UK) in 2009#
- Gordon research conference 2007: Archaea: Ecology, Metabolism, and Molecular Biology (Proctor Academy, NH, USA)*
- Molecular Biology of Archaea 2008 (St. Andrews, UK)*
- Gordon research conference 2009: Archaea: Ecology, Metabolism, and Molecular Biology (Waterville Valley Resort, NH, USA)*
- 9th Annual UK Meeting on Genetics & Molecular Mechanisms in Archaea 2010 (Birmingham, UK)#
- General Meeting of the Association for General and Applied Microbiology - VAAM, 2006 (Jena, Germany)*
- General Meeting of the Association for General and Applied Microbiology - VAAM, 2009 (Bochum, Germany)
- NWO genetics platform meeting 2006 (Lunteren, NL)*
- NWO genetics platform meeting 2007 (Lunteren, NL)#
- NWO protein platform meeting 2008 (Veldhoven, NL)
- NWO protein platform meeting 2009 (Veldhoven, NL)#
- Systems biology: statistical analysis of ~omics data (2006)
- Advanced course on applied genomics of industrial fermentation (2006)

General courses

- Safe handling with radioactive materials and sources (radiation expert 5B) (2006)
- Techniques for writing and presenting a scientific paper (2007)
- VLAG introductory week (2006)

Optional activities

- Biweekly PhD/PostDoc Meetings 2006-2009
- Weekly group meetings Bacterial Genetics 2006-2009
- PhD excursion USA, October 2006
- PhD excursion USA, April 2009
- Organization of PhD excursion USA, April 2009

*poster presentation
#oral presentation

□t MMVMDP □Lpu

□2□□□3□O□3□ □2□□□□O□39□□□□9nZZ91□Z2□09Da□□ □□n□ □9□mn□Occttttv3Z□□□□l □□□□□□□13□ Z□Zv□□3v□□□3□ □v2n□k□ □□□□3□□vOZ□□3ve □Z2□□□□Z2□3m□ □O□□31□ 9□8 vn□□v3□□□9ns□□□□□O□3□ □2□r13□ Z□nce □□□3□RzP8 P8, BL8

□39nZ□□□□l 7□tt0a□ ttf□3caa□30□□□

Bj F