UCLA UCLA Previously Published Works

Title Water-soluble carotenoid proteins of cyanobacteria

Permalink https://escholarship.org/uc/item/3dm533x9

Journal Archives of Biochemistry and Biophysics, 430(1)

ISSN 0003-9861

Author Kerfeld, Cheryl A

Publication Date 2004-10-01

Peer reviewed





Available online at www.sciencedirect.com



Archives of Biochemistry and Biophysics xxx (2004) xxx-xxx

www.elsevier.com/locate/yabbi

Water-soluble carotenoid proteins of cyanobacteria

Cheryl A. Kerfeld*

Molecular Biology Institute, UCLA Box 951570, Los Angeles, CA 90095-1570, USA

Received 12 December 2003, and in revised form 15 March 2004

Abstract

In photosynthetic organisms, carotenoids function in light harvesting and in photoprotection. In cyanobacteria, there have been numerous reports of proteins that bind exclusively carotenoids. Perhaps the best characterized of these proteins are the 35 kDa water-soluble orange carotenoid proteins (OCPs). Structural, biochemical, and genomic data on the OCP and its paralogs are gradually revealing how these proteins function in photoprotection.

© 2004 Published by Elsevier Inc.

Keywords: Carotenoid; Photoprotection; Photosynthesis; Pigment-protein; Protein structure

Photosynthetic organisms both need light and need to be protected from it. Excess excitation energy, beyond that which can be used in downstream photosynthetic processes such as carbon fixation, results in direct chemical damage to components of the photosynthetic apparatus and generates harmful reactive oxygen species. Stresses such as drought, salinity, and high temperature are all known to exacerbate the sensitivity of photosynthetic organisms to light-induced damage For optimal photosynthetic function, photosynthetic organisms must respond to variation in light intensity and environmental stresses by shifting absorbed light energy between light harvesting and photoprotective functions. Carotenoids play a central role in this balance (reviewed in [1–3]).

In photosynthesis, carotenoids carry out their specialized functions as components of integral membrane, protein-pigment complexes such as the photosynthetic reaction center and the associated light-harvesting antennae. In contrast, the cyanobacterial carotenoid proteins considered here are water-soluble and do not bind chlorophyll or other pigments. Their abundance increases under high light treatment. These characteristics suggest that they function in photoprotection rather than in light harvesting.

Biochemical characterization and distribution of the **OCPs**

Orange carotenoid proteins (OCPs)¹ are readily detectable in crude, water-soluble extracts of cyanobacteria that have been isolated from natural blooms or laboratory cultures. OCPs, 35 kDa proteins that contain a non-covalently bound keto-carotenoid, 3'-hydroxyechinenone, were first identified in Arthrospira (Spirulina) maxima, Microcystis aeruginosa, and Aphanizomenon flos-aquae in 1981 [4]. Subsequently, a similar protein was purified from Synechocystis PCC 6803 [5]. The spectral features of the Syn 6803 OCP suggested it likewise contained 3'-hydroxyechinenone, however molecular weight estimation of the carotenoid by mass spectrometry indicates a mass of 819 Da, 254 Da larger than that expected for 3'-hydroxyechinenone, suggesting that this OCP contains a carotenoid glycoside [6].

The primary structure of the OCP was first elucidated by a combination of N-terminal amino acid sequencing, gene isolation, and the subsequent identification of the corresponding open reading frame in the Syn 6803 genome [5]. A survey of the genome databases reveals that highly conserved homologs of the OCP are found in all of the cyanobacteria for which genomic data are available with the exception of the *Prochlorococci* (Table 1).

^{*} Fax: 1-310-206-3914.

E-mail address: kerfeld@mbi.ucla.edu.

^{0003-9861/\$ -} see front matter © 2004 Published by Elsevier Inc. doi:10.1016/j.abb.2004.03.018

¹ Abbreviations used: OCP, orange carotenoid protein; RCP, red carotenoid protein; NTF-2, nuclear transport factor 2.

ARTICLE IN PRESS

C.A. Kerfeld | Archives of Biochemistry and Biophysics xxx (2004) xxx-xxx

Table 1					
Orange carotenoid	proteins	found	in	cyanobacteria	l genomes

A max MPPTIDTARSIPPETLAADUVPATIARFKQLSAEDOLALIWPAYLEMGKTLTIAAPGAANMQFAENTLOETRQMTPLOOT Syn 8003 MPPTIDSARGIFPNTLAADÜVPATIARFKQLSAEDOLALIWPAYLEMGKTLTIAAPGAAŠMQLAENALKEIQAMGPLOOT Nos 7120 MAITIDSAREIPPNTLQADÜVPTVUSSFSQLNAEDQLALIWPAYLEMGKTITVAAPGAAŠMQLAEDNALKEIQAMGPLOOT Nos 7120 MAITIDSAREIPPNTLQADÜVPTVUSSFSQLNAEDQLALIWPAYTEMGRSITVAAPGAAŠMQLAQGLLEQIKQMPFEAOT Syn 8012 M-FTLDKARQIFPDTLSAAÄVPAITARFKLJSAEDQLALIWPAYTEMGRSITVAAPGAAŠMQLAQGLLEQIKQMPFEAOT N punc1 MSFTIKSAQSIFPGTLVADÜVPTVUSSFSQLNAEDQLALIWPAYTEMGRSITVAAPGAAŠMQLAQGLLEQIKQMPFEAOT N punc2 MAYTIESAQAIFAETGKPSTPIGILADFNRISLEENLLLWYAYTETGRTITKAALGASÄMLUQLMEGILPNEIKQMSHEEQT A max QAMCDLANRTDTPICRTYAŠWSPNIKLGFWYELGEFMDQGLVAPIPEGYKLSANANAVLÄTIQQLBSGQQITVLRNCVUD Syn 8003 QAMCDLANRADTPICRTYAŠWSPNIKLGFWYELGEFMDQGLVAPIPEGYKLSANANAVLÄTIQQLBSGQQITVLRNAVUĎ N punc2 MAYTIESARNIFSSTQVADÄPENTAMFAELNIDDQLAFLWYAYAELGRTITPAAPGKAŇLUQLMEGIPNEIKQMSHEEQT A max QAMCDLANRADTPICRTYAŠWSPNIKLGFWYELGEFMDQGLVAPIPEGYKLSANANAVLÄTIQQLBSGQQITVLRNAVUĎ Syn 8003 QAMCDLANRADTPICRTYAŠWSPNIKLGFWYELGEFMRQGVAPIPEGYKLSKANANAVLÄTIQQLBSGQQITVLRNAVUĎ Syn 8003 QAMCDLANRADTPICRTYAŠWSPNIKLGFWYELGEFMRQGVAPIPEGYKLSKANANAVLÄTIQQLSGQQITVLRNAVUĎ N punc1 RVMYDLANRADTPICRTYAŠWSPNIKLGFWYELGEFMRQGVAPIPEGYKLSKANANAVLÄTIQQLSGQQITVLRNAVUĎ Syn 8002 KVCDLAGKINSPISARLÄVKYVIKLGFWYELGEWMAQGTVAPIPEGYKLSKANANAVLÄTIGQLEGGQQIQURDIVLN N punc2 QLMRDLARAADTPICRTYAŠWSVNKLGFWYELGEWNAQGTVAPIPACYGMSTQVKAVLÄPKVEGQQITVLRNTVVŇ MGFDTSKLGSYQRVAEPV-VPPOTASRTKVSIEGVTNATVLQYNMNNANDPDALISLFAEDGALQPPFQKPIVGKENA Syn 8003 MGFTAGKDGKRIAEPV-VPPOTASRTKVSIEGVTNATVLVNNNNNNDPDDLISLFAEDGALQPPFQKPIVGKENA N max MGFDTSKLGSYVENEPL-VPPOTASRTKVSIEGVTNATVLVNNNNNNDPDDLIKLFYEDGALQPPFQKPIVGKENA Syn 8003 MGFTAGKDGKRIAEPV-VPPOTASRTKVSIEGVTNATVLVNNNNNNDPDDLIKLFYEDGALQPPFQKPIVGKDAT MGFDDASYKVSEPVAPTEPAPPTTPAPTKYSIEGINNSTVLQYNNNNNNNDPDDLIKLFYEDGALQPPFQKPIGGALQP MGFDDSVVPDEAPEAEDPQERTEPYSTERIDVKOUTGKVQTPWFGGNVGMIAWRPLLNPEKVFFVAIDLLASPKELLNLF N monc1 MGFDPSVVPDEAPEAEDPQERTEPYSTERIDVKVGYQTPWFGGNVGMIAWRPLLMPEKKFPVAIDLLASPKELLNLF N monc2 XFYREEQQNIL				· · · · · · · · · · · · · · · · · · ·	-				
Syn 8803 M PFTIDSARGI PPNTLAADŰV PATIARFSQLNAEDQLALÏWFAYLEMGKTLTIAAPGAAŠMQLAENALKEIQAMGPLQQT Nom 7100 MAITIDSARRI FPNTLQADŽV PALTARFNQLSAEDQLALÏWFAYLEMGKTLTIAAPGAAŠMQPAGILKQI KEMTPEQT N. punci MSTIKSAQSI FPGTLVADŮV PTVLESFSQLNAEDQLALÏWFAYLEMGKTI TVAAPGAAŠMQPAGILKQI KEMTPEQT G. wol MAFTLESAQAIFAETGKPSPI PGILADFNRLSLEDRLALLWYAYTTEGKI TVAAPGAAŠMALAQPTLDEIQAMSPDQT G. wol MAFTLESAQAIFAETGKPSPI PGILADFNRLSLEDRLALLWYAYTTEGKI TVAAPGAAŠMALAQPTLDEIQAMSPDQT Syn 8002 MYTIESAQAIFAETGKPSPI PGILADFNRLSLEDRLALLWYAYTTEGKI TRAALGAASMALVEGMLDQI KQMPPAROT N. punci MYTIESAQAIFAETGKPSPI PGILADFNRLSLEDRLALLWYAYTTEGKI TRAALGAASMALQLMEGIPNEIKQUSED Syn 8003 QAMCDLANRTDTPICKTYAŠWSPNIKLGFWYELGRFMDQGLVAPI PEGYKLSANANAILYTIQGIDFGQQI TVLKNAVUĎ Syn 8003 QAMCDLANRADTPLCKTYAŠWSPNIKLGFWYELGRFMDQGLVAPI PEGYKLSANANAVLÄTI QCLESGQQI TVLKNAVUĎ Nom 7120 QVMCDLANRADTPLCKTYAŠWSPNIKLGFWYELGELMEQĞFVAPI PEGYKLSANANAVLÄLLKELGQGQI TVLKNAVUĎ Syn 8003 QAMCDLANRADTPLCKYAŠFTVNIKLGFWYELGELMEQĞFVAPI PEGYKLSANANAVLÄLLKELGQGQI TVLKNAVUĎ Syn 8002 QVMCDLANRADTPLCKYYAŠFTVNIKLGFWYELGELMEQĞFVAPI PEGYKLSANANAVLÄLLKELGQGQI TVLKNAVUĎ Syn 8002 QVMCDLANRADTPLCKSYAŠFTVNIKLGFWYELGELMEQĞFVAPI PEGYKLSANANSVLÄLKVEQGQQI TLLKNTVVŠ Syn 8002 KVMCDLARADTPLSKSYAŠFTVNIKLGFWYELGELMEQĞFVAPI PEGYKLSANANSVLÄLKVEQGQQI TLLKNTVVĎ Syn 8002 KVMCDLARADTPISKSYGÝŠVNTKLGFWYELGELMEQĞVAPI PEGYKLSANANSVLÄLKVEQGQQI TLLKNTVVĎ Syn 8003 MGFTAGKDC	A. max	MPFTIDTARSI	FPETLAADV	VPATIARFKQ	LSAEDQLALI	WFAYLEMGKŢI	TIAAPGAANM	QFAENTLQĘIR	QMTPLQQT
Nos 7120 MAITIDSARRIF PNTLQADÄV PALTAR FNQLSAEDQLAWTWFAFLEMGKTITVAAPGAASMQFAEGILKQIKEMTFEEQT N. punci 1MSFTIKSAQSIF POTLVAD ÜV PTVVES FSQLNAEDQLALÜWFAYTEMGRITVAAPGAASMVLAQGLEGIKQMFFEQT Syn 8102 M-FTLDKARQIF PDTLSAARVPAITARFKLISAEDQLALÜWFAYTEMGRITVAAPGAASMALAQPTLDEIQAMSFDEQT G. vol MAFTLESAQAIFAETGKPS PIFOILADFNRLSLEDRLALLWYAYTETGRITRALGAASMALAQPTLDEIQAMSFDEQT N. punci 2MYTTIESARNIFSSTQVADÄV PATTAMFAELNIDDQLAFÜWAYTETGRITRALGAASMALVEGMLDQIKQMPFAEQT A. max QAMCDLANRTDTPICCRTYASWSPNIKLGFWYLLGRFMDQGLVAPIPEGYKLSANANATLÄTIQGLESGQITVLRNAVDÖ Nos 7120 QVMCDLANRTDTPICCRTYASWSPNIKLGFWYLGBLMEQĞFVAPIPAGYQLSANANAVLÄTIQGLESGQITVLRNAVDÖ Nos 7120 QVMCDLANRTDTPICCRTYASWSPNIKLGFWYLGBLMEQĞAVAPIPAGYQLSANANAVLÄTIQGLESGQQITVLRNAVDÖ Nos 7120 QVMCDLANRTDTPICCRTYASWSPNIKLGFWYLGBUNGQĞVAPIPAGYQLSANANAVLÄTIQGLESGQQITVLRNAVDÖ Nos 7120 QVMCDLANRTDTPICCRTYASWSPNIKLGFWYLGBUNGQĞVAPIPAGYQLSANANAVLÄTIQGLESGQQITVLRNAVDÖ Nos 7120 QVMCDLANRTDTPICCRTYASWSPNIKLGFWYLGGUMAQĞIVAPIPAGYQLSANANAVLÄTIQGLESGQQITVLRNAVDÖ Nos 7120 QVMCDLANRTDTPICRTYASWSPNIKLGFWYLGGUMAQĞIVAPIPAGYQLSANANAVLÄTIQALSGQQITVLRNAVDÖ Nos 7120 QVMCDLANRTDTPICRTYASWSVNKLGFWQLGBUMAQĞIVAPIPAGYQLSANANSVLÄAVKVEQQQQITLLRNTVVS Sym 8023 KVMPDLANRADTPISRSYAYFSVNKLGFWQLGBUMAQĞIVAPIPAGYQKANSVLÄAVKVEQQQQITLLRNTVVS Nos 7120 QVMCDLANRTDTPISRSYAYFSVNKLGFWQLGBUMAYĞIVAPIPANYQMSTDAQLLFESIKNLDGGQQQVLDIVLN N. punc 2 QLMRDLASNADTPISRSYAYFSVNKLGFWQLGBUMKEĞIVAPMEVGYQMSTDAQLLFESIKNLDGGQLQPPQRPIVGKENA N. punc 2 QLMRDLASNADTPISRSYAYFSVNKLGFWQLGBUMSTVLQYMDNLNANDFDLILSLFAEDGALQPPPQRPIVGKENA N. punc 1 MGFDDAAKLDGYRVAEPL - VPPQDTASRTKVSIEGVTNATVLQYMDNLNANDFDLILSLFAEDGALQPPPQRPIVGKENA N. punc 1 MGFDDAAKLDGYRVAEPL - VAFKDISQVQVITEGINNSTVLXYMNNLNANDFDLILSLFAEDGALQPPPQRPIVGRAIA N. punc 1 MGFDDAAKLDGYRVAEPL - VAFKDISQVQVITEGINNSTVLXYMNNLANDFDLILSLFAEDGALQPPPQRPIVGQDAI SM 8003 GREFFEECQNLKLIPERGVEPPAPERTEDGYTQIKVTGKVQTPWFGANVGMNIAMRFLNPGAVALFFNNGALQPPPQKPIVGQDAI N. punc 1 MGFDDSVLDPAASEPPAPERTEDGYTQIKVTGKVQTPWFGANVGMNIAMRFLNPGKVPFVAIDLLASPKELLNLF N. punc 1 MYREEQQLLKLIPEGYBFDAGFYQKVTGKVQTPWFGANVGMNIAMRFLNPGKFPVAIDLLASPKELLN	Syn 6803	MPFTIDSARGI	FPNTLAAD	VPATIARFSQ	LNAEDQLALĨ	WFAYLEMGKŢL	тіаардааё́м	QLAENALKĘIÇ	AMGPLQQŤ
N PURC 1 MSPTI KSAQŞI F PGTL VADÜV PTV VESF SQLNAEDQLALÜW FAYTEMGRŞI TVAAPGAAÑM VLAQGLLEÇI KQMPFEAQŤ Syn 8102 M- FTLDKARQI F PDTLSAAÂV FAITAR FKLLSAEDQLALÜW FAYLEMGRŞI TVAAPGAARMALAQ FTLDE I QAMSFDEQT G. viel MAPTLESAQAI FAETGKPSPÎ FGILADFNRLSLEDRLALLŴYAYTETGRȚI TRAAPGARMALAQ FTLDE I QAMSFDEQT N. PURC 2 MYNTI ESARNI FSSTQVADĂV FATTAMFABLNI DDQLAFLŴYAYBLGRTI TPAAPGKANLQLMEGI FNEI KQMSHEQT A. max QAMCDLANRTDTPI CRTYASWSPNI KLGFWYELGRFMDQGLVAPI PEGYKLSANANAL LYTIQGTD PGQQI TVLRNCVVD Syn 8603 QAMCDLANRTDTPI CRTYATWSPNI KLGFWYELGRFMDQGLVAPI PEGYKLSANANAVLÏTI QGLSGQQI TVLRNCVVD N. PURC 2 MYNTI ESARNI FSSTQVADĂV FATTAMFABLNI DDQLAFLŴYAYBLGRI PAGYQLSANANAVLÏTI KSLDQGQQI TVLRNSVD N. PURC 1 MYYDLANRADTPL CRTYATWSPNI KLGFWYELGEWMEQĞI VAPI PAGYQLSANANAVLÏTI KSLDQGQQI TVLRNSVD N. PURC 1 QVMCDLANRTDTPI CRTYATWSPNI KLGFWYELGEWMEQĞI VAPI PAGYQLSANANAVLÏTI KSLDQGQQI TVLRNSVD N. PURC 1 QVMCDLARRADTPL CRSYAFTVNI KLGFWYELGEWMEQĞI VAPI PAGYQLSANANAVLÏTI KSLDQGQQI TVLRNSVD N. PURC 1 QVMCDLARRADTPI SRSYAY MSVNVKLGFWYELGEWMAQĞI VAPI PAGYQLSANANAVLÄTI KNADSGQQI TVLRNTVVÖ G. viel XVMPDLARRADTPI SRSYAY MSVNVKLGFWYELGEFMEQĞVAPI PAGYQLSANANSVLÄVKKVEGQQI TVLRNTVVÖ A max MGFDTSKLGŞYQRVAEPV - VPPQEMSQRTKVQI EGVTNATVLNYMDNLNANDFDNLI SLFAEDGALQP FPQKPI VGKENV N. PURC 2 QLMRDLASNADTFI SRSYAY PGVNAKLGFWYELGEWMERĞI VAPMEVGYQMSTQVKAVLËAV KI DQSQQI TVLRNTVVŇ M max MGFDTSKLGŞYQRVAEPV - VPPQEMSQRTKVSI EGVTNATVLNYMDNLNANDFDNLI SLFAEDGALQP FPQRPI VGKENV N. PURC 2 MGFDAKLGYTVAEPL - VAFKDI SQUQVTI BGINNSTVLLYMNNLNANDFDELI KLFYEĞGALQP PFQRPI VGKENV N. PURC 2 MGFDDASVDPBAPEAEDPQ FERTEPVSTERI DVKVQI FYVT GYNVAMI SAMNAD FPAAVALFFI SEGALQP PFQRPI VGQDAIR SYN 8003 KGFTAGKDG KRIAEDV - VPPQEMSQRTKVSI EGVTNATVLNYMNNLNANDFDELI LFYEĞGALQP PFQKPI VGQDAIR N. PURC 2 MGFDDASVD PBAPEAEDPV PEPAPAFRTKVSI EGVTNATVLNYMNN NANDFDELI KLFYEĞGALQP PFQKPI VGQDAIR N. PURC 2 MGFDDASVD PBAPEAEDPQ FERTEPVSTERI DVKQVQT KVGVQT MYFGANVGMNI ANRFL PLNPQKVI PFYAIDLLASPKELLNLA N. PURC 2 MGFDPSVD PBAPEAEDPQ FPRTS - LSPQFTI EGVTNYLGY VMYGGNVGMNI ANRFL VANVLFANNGAL PPPQKP	Nos 7120	MAITIDSARRI	FPNTLQADÃ	VPALTARFNQ	LSAEDQLAWŤ	WFAFLEMGKŢI	туаардааรи	QFAEGILKQIK	CEMTFEEQŤ
Syn 8102 M - FTLDKARQIFPDTLSAAÄYPAITARFKLLSAEDQLALÏWFAYLEMGRTITVAAPGAARNALAQPTLDEIQAMSFDEQT G. voil MAFTLESAQAIFAETGKPST PGILADFRKLSLEDRLALIWVAYTETGRTITRAALGAASNALVEGMLDQIKQMPPAEQT N. punc 2 MNYTIESARNIFSSTQVADÄYDATTAMFAELNIDDQLAFÜWYAYAELGRTITPAAPGKANLQLMEGIFNEIKQMSHEEQT A. max QAMCDLANRADTPLCRTYASNSPNIKLGFWYRLGELMEQĞVAPIPAGYQLSANANAVLÄTIQGIESGQITVLRNCVVD Syn 8603 QAMCDLANRADTPLCRTYASNSPNIKLGFWYRLGELMEQĞVAPIPAGYQLSANANAVLÄTIQGLESGQITVLRNCVVD Nos 7120 QVMCDLANRADTPLCRTYASNSPNIKLGFWYRLGELMEQĞVAPIPAGYQLSANANAVLÄTIQGLESGQITVLRNSVVD N. punc 1 RVMYDLANRADTPLCRYASNSPNIKLGFWYRLGELMEQĞVAPIPAGYQLSANANAVLÄTLKSLDQQQITVLRNSVVD G. voil RVMPDLARRADTPLCRYASNSPNIKLGFWYQLGEWMAQĞİVAPIPAGYQLSANANAVLÄTLKSLDQQQITULRNEVVD G. voil RVMPDLARRADTPLSRYAYBVSNIKLGFWYQLGEWMAQĞİVAPIPAGYQLSANANAVLÄSNKUBGQQITLRNEVVD G. voil RVMPDLARRADTPLSRSYGYİSVNTKLGFWYQLGEWMAQĞİVAPIPANYQMSTDAQLLFEĞIKNLDGGQQIQVLRDIVLN N. punc 2 QLMRDLASNADTPISRSYGYİSVNTKLGFWYQLGEWMKEĞIVAPIPANYQMSTDAQLLFEĞIKNLDGGQQITVLRNTVVN MGFDTSKLGŞYQRVAEPV - VPPQEMSQRTKVQIEGVTNSTVLQYMDNLNANDPDNLISLFAEDGALQPPPQRPIVGKUN Syn 8603 MGFTAGKDG KRIAEPV - VPPQEMSQRTKVQIEGVTNSTVLQYMDNLNANDPDAVAVLFXBQQLQDPPQRPIVGKANL Syn 8603 MGFDARKLDĞYTRVAEPL - VAPKDISQRVQVTIEGINNSTVLWYMNNLNANDPDAVALFTSEGLQPPPQRPIVGQDAI N. punc 1 MGPDPNAPGŞYKKVSBVAPPTAPAFRTKYSIEGINNPTVLQYINMNLNANDPDAVALFTSEGLQPPPQRPIVGQDAI Syn 8102 MGPDAVALDGYTRVAEPL - VAPKDISQRVQVTIEGINNSTVLWYMNNLNANDPDAVALFTSEGLQPPPQKPIVGRDAT N. punc 1 MGPDPSVVPDRAPEREADFQFERTEPVSTERIDVGVDDTTPLKYFRAMNSDNFEAAVALFFANNGALQPPPQKPIVGRDAT A. max <u>LRFFREECQNLKLIPERGVSBPTEDGYTQIKVTGKVQTPWFGGNVGMNIAWRFLNPEKVFFVAIDLLASPKELLNL</u> - Syn 8603 LRFFREECQNLKLIPERGVSBPTEDGYTQIKVTGKVQTPWFGGNVGMNIAWRFLNPEKVFFVAIDLLASPKELLNLL - Syn 8602 LKFFKRDCQNLKLIPERGVSBPTEDGYTQVKVTGKVQTPWFGGNVGMNIAWRFLLDPQKIFFVAIDLLASPKELLNLVR N. punc 1 ÄYMREBAQGLVMKPTKGVŠETIEDGYTQVKVTGKVQTPWFGGNVGMNIAWRFLLDPQKIFFVAIDLLASPKELLNLVR N. punc 1 ÄYMREBAQGLVMKPTKGVŠETIEDGYTQVKVTGKVQTPWFGGNVGMNIAWRFLDDPCĞI'FVAIDLLASPKELLNLVR N. punc 1 ÄYMREBAQGLVMKPTKGVŠETIEDGYTQVKVTKVTGVVTPWFGGNVGNNAWRFLDDPCĞI'FVAIDLLASPKELLNLVR	N. punc 1	M SFTIKSAQSI	FPGTLVAD	VPTVVESFSQ	LNAEDQLAL	WFAYTEMGRȘI	туаардаайм	VLAQGLLEQIK	QMPFEAQŤ
G. VIOL MAFTLESAQAIFAETGKPSPIPGILADFNRLSLEDRLALLWYAYTETGRTITRAALGAAS MALVEGMLDQIKQMPPAEQT N. punc 2 MNYTIESARNIFSSTQVADA VPATTAMFAELNIDDQLAFLWYAYAELGRTITPAAPGKANLQLMEGIFNEIKQMSHEEQT A max QAMCDLANRTDTPICRTYASWSPNIKLGPWYELGRFMDQGLVAPIPGYKLSANANAVLATIQGIDFGQQITVLRNCVVD Syn 6803 QAMCDLANRTDTPICRTYASWSPNIKLGPWYELGRFMDQGLVAPIPGYKLSANANAVLATIGGLESGQQITVLRNAVVD Nos 7120 QVMCDLANHTDTPICRTYATWSPNIKLGPWYELGRFMQQGVAPIPAGYQLSANANAVLATIGGLESGQQITVLRNAVVD Syn 6803 QAMCDLANRADTPICRTYASWSPNIKLGPWYELGRFMQQGVAPIPAGYQLSANANAVLATIKSLOGQQITVLRNAVVD Nos 7120 QVMCDLANHTDTPICRTYATWSPNIKLGFWYQLGEWMQQGVAPIPAGYQLSANANAVLATIKSLOGQQITVLRNAVVD Syn 6803 QAMCDLARRADTPICRSYASYSYNKLGFWYQLGEWMQQGVAPIPAGYQLSANANAVLATIKSLOGQQITVLRNAVVD N. punc 1 RVMYDLANRADTPICRSYASYSVNKLGFWYQLGEWMQGGVAPIPAGYQLSANANAVLAGVKKVEGGQQITLLRNFVVD G. VIOI RVMPDLARRADTPISSYAYYFGVNKLGFWYQLGEWMKQGIVAPIPANYQMSTDAVLFAVKKVEGQQITVLRNIVN N. punc 2 QLMRDLASNADTPISSYAYFGVNKLGFWYQLGEWMKEGIVAPMPVGYQMSTQVKAVLEAVKKVEGQQITVLRNIVN N. punc 2 QLMRDLASNADTPISSYAYFGVNKLGFWYQLGEWMKEGIVAPMPVGYQMSTQVKAVLEAVKUDQSQQITVLRNIVN N. punc 2 QLMRDLASNADTPISSYAYFGVNKLGFWYQLGEWMKEGIVAPMPVGYQMSTQVKAVLEAVGKIDQSQQITVLRNIVN N. punc 2 QLMRDLASNADTPISSYAYFGVNAKLGFWYQLGEWMKEGIVAPMPVGYQMSTQVKAVLEAVGKIDQSQQITVLRNIVN N. punc 2 QLMRDLASNADTPISSYAYFGVNAKLGFWYQLGEWMKEGIVAPMPVGYQMSTQVKAVLEAVGKIDQSQQITVLRNIVN N. punc 2 MGFDAAKLDGYTRVAEPL - VAPKDQDTASRTKVSIEGVTNSTVLQYMDNLNANDFDLILLFFTSDGALQPPFQRPIVGKENA N. punc 1 MGFDPNAPGSYKKVSEPVAPFTAPAFRTKVSIEGINNSTVLDYNNNUNANDFDAAVALFTSGGLQPPFQRPIVGQDAI G. VIOI MGFDPSVVPDEAPEAEDFYPTPVDQREEILIPGVNQTILGYNNTNNJNANDFDAAVALFSEGALQPFPQKPIVGRAAI N. punc 2 MGFDDSLADKKQAQ-INFKFPRTS LSPQFTIEGVTBFTVLKYIEAMNADNFEAAVALFSDGALQPFPQKPIVGRAAI A. max LRFFREECQNLKLIPERGVSEPTEDGYTQIKVTGKVQTPWFGGNVGMNIAWRFLLNPEGKIFFVAIDLLASPKELLNL- JW M5012 LRFFREECQNLKLIPERGISEPVEDGYTQVKVTGKVQTPWFGGNVGMNIAWRFLLNPEGKIFFVAIDLLASPKELLNL- JW M5012 LFFFREECQNLKLIPERGISEPVEDGYTQVKVTGKVQTPWFGAVGMNIAWRFLLNPEGKIFFVAIDLLASPKELLNLVR N. punc 1 AYMREEAQGLVMKPIEGITEDGYTQKVTGKVQTPWFGAVGMNIAWRFLLDPQGKIFFVAIDLLASPKELLNLTR	Syn 8102	M-FTLDKARQI	FPDTLSAAÃ	VPAITARFKL	LSAEDQLAL ¹	WFAYLEMGRTI	TVAAPGAAŔM	ALAQPTLDEIQ	AMSFDEQŤ
N punc 2 N punc 2 N punc 2 N punc 2 N punc 2 N punc 2 N punc 2 N punc 2 N punc 2 N punc 2 N punc 2 N punc 2 N punc 2 N punc 2 N punc 2 N punc 2 N punc 2 N punc 2 N punc 2 N punc 3 N punc 4 N punc 4 N punc 4 N punc 4 N punc 4 N punc 4 N punc 5 N punc 4 N punc 5 N punc 4 N punc 5 N punc 5 N punc 4 N punc 5 N punc 5 N punc 6 N punc 6 N punc 6 N punc 7 N punc	G. viol	MAFTLESAQAI	FAETGKPSP	PGILADFNR	LSLEDRLALL	₩ŶAYTETGRŢI	TRAALGAASM	ALVEGMLDQIK	QMPPAEQŢ
A max QAMCDLANRTDTPICRTYASWSPNIKLGFWYELGRFMDQGLVAPIPEGYKLSANANAILVTIQGIDFGQQITVLRNCVVD Syn 6803 QAMCDLANRTDTPICRTYASWSPNIKLGFWYLGELMEQĞFVAPIPAGYQLSANANAVLÄTIQGLESGQQITVLRNAVVD Nos 7120 QVMCDLANRTDTPICRTYASWSPNIKLGFWYLGELMEQĞFVAPIPAGYQLSANANAVLÄTIQGLESGQQITVLRNAVVD Nos 7120 QVMCDLANRTDTPICRTYASWSPNIKLGFWYQLGEWMAQĞIVAPIPAGYQLSANANAVLÄTIKLSDQQQQITVLRNAVVD Syn 8102 KVMCDLAGKINSPISARYAYWSVNVKLGFWYELGEPMEQĞKVAPIPAGYQLSANANAVLÄTIKNDSQQTITLRNFVVD G. vol RVMTDLARRADTPISRSYGYFSVNTKLGFWYLGEBWRQĞKVAPIPAGYQLSANANSVLÄZVKKVEQGQQITVLRNTVVS Syn 8102 KVMCDLAGKINSPISARYAYWSVNVKLCFWYELGEPMEQĞKVAPIPAGYQKSTQVKAVLËSVKKVEQGQQITVLRNTVVS M. punc 2 QLMRDLASNADTPISRSYGYFSVNTKLGFWYQLGEWMKEĞIVAPMPVGYQMSTDAQLLFEĞIKNLDGGQQIQVLRDIVLN N. punc 2 QLMRDLASNADTPISRSYGYFSVNTKLGFWYQLGEWMKEĞIVAPMPVGYQMSTDAQLLFEĞIKNLDGGQQIQVLRDIVLN N. punc 2 QLMRDLASNADTPISRSYGYFSVNTKLGFWYQIEGVTNSTVLQYMDNLNANDFDNLISLFAEDGALQPPFQKPIVGKENA MGFDTSKLGŞYQRVAEPV - VPPQEMSQRTKVQIEGVTNSTVLQYMDNLNANDFDNLISLFAEDGALQPPFQKPIVGKENA N. punc 1 MGFDPAKLGŞYKKVSEPVAPFTDPQTASRTKVSIEGVTNSTVLQYMDNLNANDFDLIELFFTSDGALQPFPQKPIVGKENA Nos 7120 MGFDDAKLDGYTRVAEPL - VAPKDISQRVQVITEGINNSTVLŇYMNNLNANDFDAAVALFTŠEGGLQPPFQRPIVGQDAIF Syn 8102 MGFDPSVPDEAPEAEDFPERTEPVSTERIDVKGVDD PTPLŽYTRANDLANDFDQLIDIFLNGALQPPPQKPIVGRDAI G. vol MGFDPSVVPDEAPEAEDFQFRETEPVSTERIDVKGVDD PTPLŽYTEANNSDNFBAAVALFEPËGALQPPFQKPIVGRDAI M. punc 2 MGFDDSLADKKQQQ - INFKFPRTS - LSPQFTIEGVTEPTVLKYIEANNSDNFBAAVALFPSEGALQPPFQKPIVGREAT M. punc 2 MGFDDSLADKKQQQ - INFKFPAEDGFTQIKVTGKVQTPWFGGNVGMNIAWRFLLNPEGKIFPVAIDLLASPKELLNL - Syn 6803 LFFFREECQNLKLIPERGVSEPTEDGYTQIKVTGKVQTPWFGGNVGMNIAWRFLLNPEGKIFPVAIDLLASPKELLNL - Syn 6803 LFFFREEQQLLKMIPERGISËVEDGTTQVKVTGKVQTPWFGGNVGMNIAWRFLLDPEGKIFPVAIDLLASPKELLNLAF Syn 8102 LKFFFREQQLLKLIPERGISËVEDGTTQVKVTGKVQTPWFGAVGMNAWRFLLDPEGKIFFVAIDLLASPKELLNLR Syn 8102 LKFFFREQQLLKKIPERGISEVEDGTTQVKVTGKVQTPWFGAVGMNIAWRFLLDPEGKIFFVAIDLLASPKELLNLR Syn 8102 LKFFFREQQLLKKIPERGISËVEDGTTQVKVTGKVQTPWFGAVGMNIAWRFLLDPEGKIFFVAIDLLASPKELLNLR Syn 8102 LKFFFKBUCQNLKIPEGIFFVFGKFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF	N. punc 2	¹ MNYTIESARNI	FSSTQVADA	VPATTAMFAE	LNIDDQLAF	WYAYAELGRTI	TPAAPGKAÑL	QLMEGIFNEIF	QMSHEEQ ^{bo}
A. max QAMCDLANRTDTPICRTYASWSPNIKLGFWYELGRFMDQGLVAPIPEGYKLSANANAILVTIQGIDPGQQITVLRNCVVD Syn 6803 QAMCDLANRADTPLCRTYASWSPNIKLGFWYELGELMEQÖFVAPIPAGYQLSANANAVLÄTIQGLESGQQITVLRNCVVD Nos 7120 QVMCDLANHTDTPICRTYASWSPNIKLGFWYELGELMEQÖFVAPIPAGYQLSANANAVLÄTIQGLESGQQITVLRNAVD N. punc1 RVMYDLANRADTPLCRYAS SPTVNIKLGFWYQLGEWMAQĞIVAPIPAGYQLSANANAVLËTLKSLDQGQQITVLRSSVD Syn 8102 KVMCDLAGKINSPISARYAY WSVNVKLCFWYELGEFMRQĞKVAPIPAGYQLSANANSVLËAVKKVEQGQQITLLRNTVVŠ Syn 8102 KVMCDLAGKINSPISARYAY WSVNVKLCFWYELGEFMRQĞKVAPIPAGYQKSPAADVLËAVKKVEQGQQITVLRNIV Syn 8102 KVMCDLAGKINSPISARYAY WSVNKLCFWYELGEFMRQĞKVAPIPAYQMSTQVKAVLËAVKKVEQGQQITVLRNIV N. punc2 QLMRDLASNADTPISRSYGYF SVNTKLGFWYQLGEWMAQĞIVAPIPANYQMSTQVKAVLËAVKVEQGQQITVLRNIVN A. max MGFDTSKLGŞYQRVAEPV - VPPQEMSQRTKYQIEGVTNSTVLQYMDNLNANDFDNLISLFAEDGALQPPFQKPIVGKENA Syn 8603 MGFTAGKDG - KRIAEPV - VPPQDTASRTKVSIEGVINSTVLQYMDNLNANDFDLILFF SDGALQPFPQKPIVGKENA Nos 7120 MGFDAKLDGYTRVAEPL - VAPKDISQRVQVTIEGINNSTVLÑYMNNLNANDFDAVALFT SEGGLQPPFQKPIVGKENA N. punc1 MGFDPSVVPDEAPEAEDFQFERTEPVSTERIDVKGVDPTPLRYFEAMNSDNFEAAVALFFPEGALQPFPQKPIVGRAAT G. viol MGFDPSVVPDEAPEAEDFQFERTEPVSTERIDVKGVDPTPLRYFEAMNSDNFEAAVALFPEGALQPFPQKPIVGRAAT N. punc2 MGFDSLADKKQAQ - INFKFPRTS - LSPQFTIEGVTGYVCYTVKVGKVQTPWFGGNVGMNIAWRFLLNPEGKŮPFVAIDLLASPKELLNL - Syn 6803 LRFFREECQNLKLIPERGVEPTEDGYTQIKVTGKVQTPWFGGNVGMNIAWRFLLNPEGKŮPFVAIDLLASPKELLNL - Syn 6803 LFFREECQNLKLIPERGVEPTEDGYTQIKVTGKVQTPWFGGNVGMNIAWRFLLNPEGKŮPFVAIDLLASPKELLNL - Syn 6803 LFFFREECQNLKLIPERGYAE PÅDDGYTQVKVTGKVQTPWFGGNVGMNIAWRFLLNPQGKIFFVAIDLLASPKELLNL S Syn 120 LKFFREECQNLKLIPERGYAE PÅDDGYTQVKVTGKVQTPWFGGNVGMNIAWRFLLDPQGKIFFVAIDLLASPKELLNL S Syn 6803 LFFFREECQNLKLIPERGYAE PÅDDGYTQVKVTGKVQTPWFGGAVGMNIAWRFLLDPQGKIFFVAIDLLASPKELLNL S Syn 6102 LKFFREECQNLKLIPERGYAE PÅDDGYTQVKVTGKVQTPWFGGAVGMNIAWRFLLDPQGKIFFVAIDLLASPKELLNL S Syn 6102 LKFFREECQNLKLIPERGYAE PÅDDGYTQVKVTGKVQTPWFGGAVGMNIAWRFLLDPQGYIFVAIDLLASPKELLNL S Syn 6102 LKFFREECQNLKLIPERGYAE PÅDDGYTQVKVTGKVQTPWFGGAVGMNIAWRFLLDPQGYIFFVAIDLLASPKELLNL S Syn 6102 LKFFREEQQNLKLKPYGYEFTEFTVFDYFFFTFFFFFFNDFLANGAVGNI		0 1		20		40	60		
Syn 6003 QAMCDLANRADTPLCRTYAS WSPNIKLGFWYRLGELMEQ GFVAPIPAGYQLSANANAVLÄTLGGLESGQQITVLRNAVUD Syn 6003 QAMCDLANRADTPLCRTYAS WSPNIKLGFWYRLGELMEQ GFVAPIPAGYQLSANANAVLËTLGGLESGQQITVLRNAVUD Nos 7120 QVMCDLANRADTPLCRTYAS WSPNIKLGFWYQLGEWMQG GVAPIPAGYQLSANANAVLËTLSSLDQGQITVLRNAVUD N. punc1 RVMYDLANRADTPLCRSYAS PTVNIKLGFWYQLGEWMQG GVAPIPAGYQLSANANAVLË ARNADGQQITLLRNPVD G. viol RVMFDLARRADTPISRSYGY SVNVKLCFWYELGEFMRQ GKVAPIPQGYLSANANSVLËAVKKVEQGQQITLLRNPVD G. viol RVMFDLARRADTPISRSYGY SVNVKLCFWYELGEFMRQ GKVAPIPQGYLASINANSVLËAVKVEQGQQITVLRNVV N. punc2 QLMRDLASNADTPISRSYGY SVNVKLCFWYELGEFMRQ GKVAPIPANYQMSTDAQLLFESIKNLDGGQQIQVLRDIVLN N. punc2 QLMRDLASNADTPISRSYGY SVNVKLGFWYQLGEWMKEGIVAPIPANYQMSTDAQLLFESIKNLDGGQQITVLRNVV 1000 A. max MGFDTSKLGSYQRVAEPV - VPPQEMSQRTKVQIEGVTNSTVLQYMDNLNANDFDNLISLFAEDGALQPFPQ KPIVGKENA Syn 6803 MGFTAGKDG KRIAEPV - VPPQEMSQRTKVQIEGVTNSTVLQYMDNLNANDFDLISLFAEDGALQPFPQ KPIVGKENA Syn 6803 MGFTAGKDG KRIAEPV - VPPQEMSQRTKVQIEGVTNSTVLQYMDNLNANDFDLISLFAEDGALQPFPQ KPIVGKENA Syn 6803 MGFDAKLDG YTRVAEPL - VAPKDISQRVQVTIEGINNSTVLNYMNNLNANDFDLISLFAEDGALQPFPQ KPIVGKENA Syn 6803 MGFDAKLDG KRIAEPV - VPPQEMSQRTKVSIEGVNATVLNYMNNLNANDFDLISLFAEDGALQPFPQ KPIVGKENA Nos 7120 MGFDDAKLDG VAPKDISQRVQVTIEGINNSTVLNYMNNLNANDFDLISLFAEDGALQPFPQ KPIVGKENA Syn 6803 MGFDASVVDDEAPEAEDFQFERTEPVSTERIDVKGVDTPLFRVLNYMNNLNANDFDLIKLFVEDGALQPFPQ KPIVGRDAT G. viol MGFDPSVVD, EAPEAEDFQFERTEPVSTERIDVKGVDTPWFGNVGMNIANDFDAAVALFTSEGGLQPFPQ KPIVGRDAT MS 7120 MGFDPSULADKKQAQ - INFKPRTS - LSPQFTIEGVTEPTVLKYIEAMNADNFEAAVALFANNGALQPFPQ KPIVGRAT MS 7120 MFFREECQNLKLIPERGVEPAEDGYTQIKVTGKVQTPWFGNVGMNIAWRFLLNPEGKTFFVAIDLLASPKELLNL- Syn 6803 LFFFREECQNLKLIPERGVEPAEDGYTQVKVTGKVQTPWFGNVGMNIAWRFLLNPEGKTFFVAIDLLASPKELLNLTR Syn 8102 LKFFFREECQNLKLIPERGVEPAEDGYTQVKVTGKVQTPWFGRSVGMNAWRFLLNPEGKTFFVAIDLLASPKELLNLTR Syn 8102 LKFFFREEQQLLKLIPERGVEPAEDGYTQVKVTGKVQTPWFGRSVGMNIAWRFLLNPEGKIFFVAIDLLASPKELLNLTR Syn 8102 LKFFFREEQQLLKLIPEGYTEPTEGGFNQIKVTGKVQTPWFGRSVGMNIAWRFLLNPEGKIFFVAIDLLASPKELLNLTR Syn 8102 LKFFFREEQQLLKKPFKYSTEFIEGGTYQVKVTGKVQTPWFGRSVG	A may	OAMODIANDTD		WEDNIKICEW	VELOPEMDOC	TWADTDECVEL	CANANATIVT	TOGTOPCOOT	WIDNOWUD
Syn 6803 QAMCDLANRADTPLCRTYASWSPNIKLGFWYRLGELMEQGFVAPIPAGYQLSANANAVLÄTIQGLESGQQITVLENAVVD Nos 7120 QVMCDLANHTDTPLCRSYA YFYNIKLGFWYQLGEWMEQGAVAPIPAGYQLSANANAVLËTLKSLDQGQITVLENSVVÖ N. punc1 RVMYDLANRADTPLCRSYA YFYNIKLGFWYQLGEWMAQGYVAPIPAGYQLSANANAVLËTLKSLDQGQQITULRNTVV Syn 8102 KVMCDLAGKINSPISARYA YWSVNVKLCFWYELGEFMRQ KVAPIPQGYR, SANANSVLËAVKKVEQGQQITLLRNFVVÖ G. viol RVMFDLARRADTPISRSYGYFSVNTKLGFWYQLAEWMAQGYVAPIPANYQMSTDAQLLFB'SIKNLDGQQQQIVLKDIVLN N. punc2 QLMRDLASNADTPISRSYGYFGVNAKLGFWYQLAEWMAQGYVAPIPANYQMSTDAQLLFB'SIKNLDGQQQITVLRNTVV M. punc2 QLMRDLASNADTPISRSYAYFGVNAKLGFWYQLGEWMKEGIVAPMPVGYQMSTQVKAVLËAVQKIDQSQQITVLRNTVV M. punc2 QLMRDLASNADTPISRSYAYFGVNAKLGFWYQLGEWMKEGIVAPMPVGYQMSTQVKAVLËAVQKIDQSQQITVLRNTVV M. punc2 QLMRDLASNADTPISRSYAYFGVNAKLGFWYQLGEWMKEGIVAPMPVGYQMSTQVKAVLËAVQKIDQSQQITVLRNTVV M. punc2 QLMRDLASNADTPISRSYAYFGVNAKLGFWQLGEWMKEGIVAPMPVGYQMSTQVKAVLËAVQKIDQSQQITVLRNTVV M. punc2 QLMRDLASNADTPISRSYAYFGVNAKLGFWQLGEWMKEGIVAPMPVGYQMSTQVKAVLËAVQKIDQSQQITVLRNTVV M. punc2 QLMRDLASNADTPISRSYAYFGVNAKLGFWQQLGEWMKEGIVAPMPVGYQMSTQVKAVLËAVQKIDQSQQITVLRNTVV M. punc2 QLMRDLASNADTPISRSYAYFGVNAKLGFWQLGEWMKEGIVAPMPVGYQMSTQVKAVLËAVQKIDQSQQITVLRNTV M. punc2 QLMRDLASNADTPISRYAYFGVNAKUGFWQLGEWMKEGIVAPMPVGYQMSTQVKAVLËAVQKIDQSQQITVLRNTVV M. punc2 MGFDASKLGSYQVAEPV-VPPQEMSQRTKVQIEGVTNSTVLQYMDNLNANDFDLLIKLFVEDGALQPPFQRPIVGKENN N. punc1 MGFDANKLGSYVVDEAPHTVDQREBILIPGVNQVTIEGINNSTVLQYMDNLNANDFDLILKLFVEDGALQPFQRPIVGQDAIR Syn 8102 MGFDASKKVSEPVAPPTAPAFRTKVSIEGINNPTVLGYILSYMQLLNANDFDQLILJFLNDGALQPFQRPIVGRATI N. punc2 MGFDPSLADKKQQ-INFKFPRTSLSPQFTIEGVTEPTVLKYIEAMNSDNFEAAVALFFANNGALQPFQKPIVGREAT 100 A. max LRFFREECQNLKLIPERGVEPTDGYQIKVTGKVQTPWFGGNVGMNIAWRFLLNPEGKIFFVAIDLLASPKELLNL 100 A. max LRFFREECQNLKLIPERGVEPTDGYQVKVTGKVQTPWFGGAVGMNIAWRFLLDPQGKIFFVAIDLLASPKELLNLVR N. punc2 1ÅYMREEQQLLKMIPERGISËPVEDGYTQVKVTGKVQTPWFGGAVGMNIAWRFLLDPQKIFFVAIDLLASPKELLNLVR N. punc2 1ÅYMREEQQLLKMIPERGISËFVEDGYTQVKVTGKVQTPWFGGAVGMNIAWRFLLDPQKIFFVAIDLLASPKELLNLVR N. punc2 1ÅYMREEQQLLKMPTGQTMKFTGVEFTGGFNQIKVTFFFFGGNVAMNIAWRFLDPQGQIYFVAIDLLASPAELLKLG	A. max	QANCD LANK I D		MPENIKIGEM		PAMELEFGIÚP	JANANAL DVI	IŐGIDEGŐŐII	
Nos 7120 QVMCDLANHTDTPICRTYAŤWSPNIKLGFWNQLGEWMEQĞAVAPIPAGYQLSANANAVLËTLKSLDQGQQITVLRSSVU N. punc 1 RVMYDLANRADTPLCRSYAŠFTVNIKLGFWYQLGEWMAQĞIVAPIPAGYQLSANANAVLËATRNADSGQQITILRNTVVŠ Syn 8102 KVMCDLAGKINSPISARYAŸWSVNVKLCFWYELGEPMEQĞKVAPIPQGYRLSANANSVLËAVKKUEQGQQITLENPVVĎ G. viol ŘVMFDLARRADTPISRSYGYËSVNKLGFWYQLAEWMAQĞÎVAPIPAQYQMSTDAQLLFEŠIKNLDGGQIQVLRDIVLN N. punc 2 QLMRDLASNADTPISRSYGYËSVNKLGFWYQLAEWMAQĞÎVAPIPANYQMSTDAQLLFEŠIKNLDGQQITVLRNTVVŇ MGFDTSKLGŞYQRVAEPV-VPPQEMSQRTKVQIEGVTNSTVLQYMDNLNANDFDNLISLFAEDGALQPPFQKPIVGKENA N. punc 2 QLMRDLASNADTPISRSYAYËPGVNAKLGFWQLGEWMKEĞIVAPMPVGYQMSTQVKAVLËAVQKIDQSQITVLRNTVVŇ MGFDTSKLGŞYQRVAEPV-VPPQEMSQRTKVQIEGVTNSTVLQYMDNLNANDFDNLISLFAEDGALQPPFQKPIVGKENA N. punc 2 QLMRDLASNADTPISRSYAYËPGVNAKLGFWQLGEVMKEĞIVAPMPVGYQMSTQVKAVLËAVQKIDQSQITVLRNTVVŇ MGFDTSKLGŞYQRVAEPV-VPPQEMSQRTKVQIEGVTNSTVLQYMDNLNANDFDNLISLFAEDGALQPPFQKPIVGKENA N. punc 2 MGFDAKLDGYTRVAEPL-VAPŘĎISQRVQVTIEGINNSTVLŇYMNNLNANDFDLIKLFVEĎGALQPPFQKPIVGKENA N. punc 1 MGFDPNAPGSYKKVSEPVAPŤTAPAFRTKVSIEGVNATČUNYMNNLNANDFDLIKLFVEĎGALQPPFQKPIVGKENA N. punc 2 MGFDPSVVPDEAPEAEDFQEITĚPVDQREEILIPGVLQŤILSYMQLLNANDFDQLIDLĚČINDGALQPPFQKPIVGRAIŤ G. viol MGFDPSVVPDEAPEAEDFQEITĚPVDQREEILIPGVNQVTPVĚGGNVGMNIAWRFLLNPENKVFFVAIDLLASPKELLNL- Syn 6803 ŮRFFREECQNLKLIPERGVEĚPAEDGFTQIKVTGKVQTPWFGGNVGMNIAWRFLLNPEKVFFVAIDLLASPKELLNL- Syn 6020 ŮRFFREECQNLKLIPERGVEĚPAEDGFTQIKVTGKVQTPWFGAXOGMNAWRFLLDPGKIFFVAIDLLASPKELLNLA N. punc 2 MGPDPSLADKKQAQ - INFKĚPRTS - LSPQFTIEGVTPFŘGGNVGMNIAWRFLLDPGKIFFVAIDLLASPKELLNLA N. punc 2 MGPDSLADKKQAQ - LNKLPERGVEĚPAEDGFTQIKVTGKVQTPWFGAXOGMNIAWRFLLDPGKIFFVAIDLLASPKELLNLA N. punc 1 ÅYMREECQNLKLIPERGVEFVEDGYQVKVTGKVQTPWFGAXOGNNIAWRFLLDPGKĬFFVAIDLLASPKELLNLA N. punc 2 AMGPDSLADKKQQ FYKJGEFTEGGFTQIKVTGKVQTPWFGAXOGNNIAWRFLLDPGKĬFFVAIDLLASPKELLNLA N. punc 2 AYMREECQNLKKITGGITEVĚDGSTVQIKVTGKVQTPWFGAXOGNNIAWRFLLDPGKĬFFVAIDLLASPKELLNLA N. punc 2 AYMREECQQLKKMIPERGISĚPTEGGFTQIKVTGKVQTPWFGAXOGNNIAWRFLLDPGKIFFVAIDLLASPKELLNLA N. punc 2 AYMREECQQLKKKTGVZŠEŤIEDGYTQVKVTGKVQTPWFGAXOGNNIAWRFLLDEDŘÍ FYVAIDLLASPKELLNLŘ	Syn 6803	QAMCDLANRAD	TPLCRTYAS	WSPNIKLGFW	YRLGELMEQG	FVAPIPAGYQL	SANANAVLAT	IQGLESGQQII	VLRNAVVD
N. punc 1 RVMYDLANRADTPLCRSYAÖFTVNIKLGFWYQLGEWMAQÖIVAPIPEGYKLSPKAADVLËAIRNADSGQQITILRNTVVÖ Syn 8102 KVMCDLAGKINSPISARYAŸWSVNVKLCFWYELGEFMRQÖKVAPIPQGYRLSANANSVLËAVKKVEQGQQITLLRNFVVĎ G. viol ŘVMPDLARRADTPISRSYGYŠSVNTKLGFWYQLAEWMAQGŤVAPIPANYQMSTDAQLLFEŠIKNLDGGQQIQVLRDIVLN N. punc 2 QLMRDLASNADTPISRSYGYČGVNAKLGFWYQLGEWMEGIVAPPPQGYQMSTQVKAVLËAVQKIDQSQQITVLRNTVVŇ A. max MGFDTSKLGSYQRVAEPV - VPPQEMSQRTKVQIEGVTNSTVLQYMDNLNANDFDNLISLFAEDGALQPPFQKPIVGKENA Syn 6803 MGFTAGKDG - KRIAEPV - VPPQEMSQRTKVQIEGVTNSTVLQYMDNLNANDFDNLISLFAEDGALQPPFQKPIVGKENA Nos 7120 MGFDAAKLDGYTRVAEPL - VAPKDISQRVQVTIEGINNSTVLÑYMNNLNANDFDALIKLFYEĎGALQPPFQRPIVGKENA N. punc 1 MGFDPAARLDGYTRVAEPL - VAPKDISQRVQVTIEGINNSTVLÑYMNNLNANDFDAAVALFTŠEGGLQPPFQRPIVGKDAI Syn 6803 MGFTAGKDG - KRIAEPV - VPPQEMSQRTKVSIEGVTNATVLNYMNLNANDFDAAVALFTŠEGGLQPPFQRPIVGKENA Nos 7120 MGFDAAKLDGYTRVAEPL - VAPKDISQRVQVTIEGINNSTVLÑYMNNLNANDFDAAVALFTŠEGGLQPPFQRPIVGQDAIR Syn 8102 MGFDPSVVPDEAPEAEDFQFERŤEPVSTERIDVKGVDDPTPLŘYFEAMNSDNFEAAVALFPEĞGALQPPFQRPIVGRAAI G. viol MGFDPSVVPDEAPEAEDFQFERŤEPVSTERIDVKGVDDPTPLŘYFEAMNSDNFEAAVALFPEĞGALQPPFQKPIVGREAI N. punc 2 MGFDPSLADKKQAQ - INFKFPRTS LSPQFTIEGVTEPŤVLKYIEAMNADNFEAAVALFEPEĞALQPPFQKPIVGREAI N. punc 2 100 FFFREECQNLKLIPERGVSEPTEDGYTQIKVTGKVQTPWFĞGNVGMNIAWRFLLNPEGKIPFVAIDLLASPKELLNL Syn 6803 ÜRFFREECQNLKLIPERGVAEPŤEPAEDGFTQIKVTGKVQTPWFĞGNVGMNIAWRFLLNPEGKIPFVAIDLLASPKELLNLR N. punc 1 ÅYMREECQLLKMIPERGISEPTEDGYTQVKVTGKVQTPWFĞAVGMNIAWRFLLNPEGKIPFVAIDLLASPKELLNLR N. punc 1 ÅYMREECQNLKLIPERGVAEPÄDDGYTQVKVTGKVQTPWFĞAVGMNIAWRFLLNPQGKIPFVAIDLLASPKELLNLR N. punc 1 ÅYMREECQNLKLIPERGVEFTEGGENQIKVTGKVQTPWFĞAVGMNIAWRFLLNPQGKIPFVAIDLLASPKELLNLR N. punc 1 ÅYMREECQNLKLIPEGGYGEYEGGFNQIKVTGKVQTPWFGAVGMNIAWRFLLDPQGKIPFVAIDLLASPKELLNLR N. punc 1 ÅYMREECQNLKKPTKGVĚEFTEGGFNQIKVTGKVQTPWFGAVGMNIAWRFLLDPQGKIPFVAIDLLASPKELLNLR N. punc 1 ÅYMREECQNLKKPTKGVĚEFTEGGFNQIKVTGKVQTPWFGAVVANRFLLDENDŘIYFVAIDLLASPKELLNLR N. punc 2 TAYLRDEQQLVMKPTKGVĚETEGGFNQIKVTGKVQTPWFGAVNANIAWRFLLPPGKIPFVAIDLLASPKELLNLR	Nos 7120	QVMCDLANHŢD	TPICRTYAT	WSPNIKLGFW	NQLGEWMEQG	AVAPIPAGYQL	SANANAVLET	LKSLDQGQQII	VLRSSVVD
Syn 8102 KVMCDLAGKINSPISARYAŸWSVNVKLCFWYELGEFMRQĞKVAPIPQGYRLSANANSVLÄAVKKVEQGQQITLLRNFVVĎ G. viol ŘVMFDLARRADTPISRSYGYËSVNTKLGFWYQLAEWMAQGŤVAPIPANYQMSTDAQLLFBŠIKNLDGGQQQUVLRDIVLN N. punc2 QLMRDLASNADTPISRSYAŸFGVNAKLGFWYQLGEWMKBĞIVAPMPVGYQMSTQVKAVLĚAVQKIDQSQQITVLRNTVVŇ A. max MGFDTSKLGSYQRVAEPV - VPPQEMSQRTKVQIEGVTNSTVLQYMDNLNANDFDNLISLFAEDGALQPPFQKPIVGKENA Syn 8603 MGFTAGKDG KRIAEPV - VPPQEMSQRTKVQIEGVTNSTVLQYMDNLNANDFDLIELFŤSDGALQPPFQKPIVGKENA Nos 7120 MGFDAAKLDGYTRVAEPL - VAPKDISQRVQVTIEGINNSTVLŴYMNNLNANDFDLIKLFVEĎGALQPPFQKPIVGKENV Nos 7120 MGFDAAKLDGYTRVAEPL - VAPKDISQRVQVTIEGINNSTVLŴYMNNLNANDFDLIKLFVEĎGALQPPFQKPIVGKENZ Syn 8603 MGFTAGKDG KRIAEPV - VPPQEMSQRTKVSIEGVTNATŮLNYMNLNANDFDLIKLFVEĎGALQPPFQKPIVGKENZ Nos 7120 MGFDAAKLDGYTRVAEPL - VAPKDISQRVQVTIEGINNSTVLŴYMNNLNANDFDAAVALFFŠEGGLQPPFQKPIVGKAL Syn 8102 MGFDPNDPGSYKKVSEPVAPPŤAPAFRTKVSIEGINN FVLŰGYINNMNANDFDAAVALFFŠEGGLQPPFQKPIVGRDAT G. viol MGFDPSVVPDEAPEAEDFQFERŤEVSTERIDVKGVDDPTPLŘYFEAMNSDNFEAAVALFPĚGALQPPFQKPIVGREAI N. punc2 MGFDPSLADKKQAQ - INFKFPRTS - LSPQFTIEGVTEPŤVLKYIEAMNADNFEAAVALFÂNNGALQPPFQKPIVGREAŤ A. max LRFFREECQNLKLIPERGVSEPTEDGYTQIKVTGKVQTPWFGGNVGMNIAWRFLLNPEGKĬFVAIDLLASPKELLNL Syn 6803 ÎRFFREECQNLKLIPERGVEPŤEDGYTQIKVTGKVQTPWFGGNVGMNIAWRFLNPEGKIFÝVAIDLLASPKELLNVR N. punc1 ÂYMREECQGLLKMIPERGISĚPVEDGYTQVKVTGKVQTPWFGASVGMNIAWRFLDPQGKIFÝVAIDLLASPKELLNVR Syn 8102 LKFFKRDCQNLKLPEGGYĞEPTEGGPNQIKVTGKVQTPWFGASVGMNIAWRFLDPQGKIFÝVAIDLLASPKELLNVR N. punc1 ÂYMREEAQGLLKMIPERGISĚPVEDGYTQVKVTGKVQTPWFGÅSVGMNIAWRFLDPQGKIFÝVAIDLLASPKELLNVR N. punc2 LKFFKRDCQNLKLMPGGGYĞEPTEGGPNQIKVTGKVQTPWFGÅSVGMNIAWRFLDPDGKIFÝVAIDLLASPKELLNVR N. punc2 LKFFKRDCQNLKMPIEGITEVĽPDGSKKLKVTGKVQTPWFGÅSVGMNIAWRFLDPEGKIFÝVAIDLLASPKELLNLR N. punc2 LKFFKRDCQNLKMPIEGITEVĽPDGSKKLKVTGKVQTPWFGÅSVGMNIAWRFLDPOGŮYFVAIDLLASPKELLNLR N. punc2 TAYLRDEGQLVMKFTKGVŠETIEDGYTQKKTGVVGTPVFFØŇFGRVGMNIAWRFLDPOGŮYFVAIDLLASPKELLNLRŘ	N. punc 1	RVMYDLANRAD	TPLCRSYAS	FTVNIKLGFW	YQLGEWMAQĞ	IVAPIPEGYĶL	SPKAADVLEA	IRNADSGQQIT	ILRNTVVS
G. viol ŘVMFDLARRADTPISRSYGYËSVNTKLGFWYQLAEWMAQGËVAPIPANYQMSTDAQLLFEĞIKNLDGGQQIQVLRDIVLN N. punc 2 QLMRDLASNADTPISRSYAŸFGVNAKLGFWYQLGEWMKEĞIVAPMPVGYQMSTQVKAVLËAVQKIDQSQQITVLRNTVVŇ MGFDTSKLGSYQRVAEPV - VPPQEMSQRTKVQIEGVTNSTVLQYMDNLNANDFDNLISLFAEDGALQPPFQKPIVGKENA Syn 6803 MGFTAGKDG KRIAEPV - VPPQEMSQRTKVQIEGVTNSTVLQYMDNLNANDFDTLIELFËSDGALQPPFQRPIVGKENV Nos 7120 MGFDAAKLDGYTRVAEPL - VAPKDISQRVQVTIEGINNSTVLŴMNNLNANDFDELIKLFVEDGALQPPFQRPIVGKENV Nos 7120 MGFDAAKLDGYTRVAEPL - VAPKDISQRVQVTIEGINNSTVLŴMNNLNANDFDELIKLFVEDGALQPPFQRPIVGKENV Nos 7120 MGFDAAKLDGYTRVAEPL - VAPKDISQRVQVTIEGINNSTVLŴYNNNLNANDFDELIKLFVEDGALQPPFQRPIVGKDAI Syn 8102 MGFDPNAPGSYKKVSEPVAPPTAPAFRTKVSIEGINNFTVLĜYINNMNANDFDAAVALFTŠEGGLQPPFQRPIVGQDAIR N. punc 1 MGFDPNAPGSYKKVSEPVAPPTAPAFRTKVSIEGINNFTVLÅGYINNMNANDFDAAVALFTŠEGGLQPPFQRPIVGQDAIR Syn 8102 MGFDPSVVPDEAPEAEDFQFERŤEPVSTERIDVKGVDDPTPLËŸFEAMNSDNFEAAVALFPËGALQPPFQKPIVGRDAŤ G. viol MGFDPSVVPDEAPEAEDFQFERŤEPVSTERIDVKGVDDPTPLËŸFFEAMNSDNFEAAVALFPËGALQPPFQKPIVGREAI N. punc 2 MGFDPSLADKKQAQ - INFKFPRTS - LSPQFTIEGVTEPŤVLKYIEAMNADNFEAAVALFÊNNGALQPPFQKPIVGREAI N. punc 2 LFFFREECQNLKLIPERGVEPTEDGYTQIKVTGKVQTPWFGGNVGMNIAWRFLLNPEGKĨPFVAIDLLASPKELLNL - Syn 6803 LRFFREECQNLKLIPERGVAEPÄDGFTQIKVTGKVQTPWFGGNVGMNIAWRFLLNPEGKĨPFVAIDLLASPKELLNL - N. punc 1 ÅYMREECQGLLKMIPERGIEFDDGYTQVKVTGKVQTPFFGGNVGMNIAWRFLDPQGKĨPFVAIDLLASPKELLNLVR N. punc 1 ÅYMREECQGLLKMIPERGIEFTEGGYTQIKVTGKVQTPFFGGNVGMNIAWRFLDPQGKĨPFVAIDLLASPKELLNLVR Syn 8102 LKFFKRDCQNLKLMPQGYŐËPTEGGFNQIKVTGKVQTPFFGGNVAMNIAWRFLDENDKŶIFVAIDLLASPKELLNLVR N. punc 1 ÅYMREECQGLLKKIPEGITEVŤPDGSKKLKVTGKVQTPFFGGNVGMNIAWRFLDEDDKŶIYVAIDLLASPKELLNLR N. punc 2 TAYLRDEGQGLVMKPTKGVĚFTIEDGYTQHKITGVETPŴFGGNVGMNIAWRFLLDPEGKIFÝVGIDLLASPKELLNLVŘ	Syn 8102	KVMCDLAGKIŅ	SPISARYAY	WSVNVKLCFW	¥ELGEFMRQĞ	KVAPI PQGYRL	SANANSVLEA	VKKVEQGQQII	LLRNFVVD
N. punc 2 0 LMRDLASNADTPISRSYAŸFGVNAKLGFWWQLGEWMKEĞI VAPMPVGYQMSTQVKAVLËAVQKIDQSQQI TVLRNTVVŇ A. max MGFDTSKLGSYQRVAEPV - VPPQEMSQRTKVQIEGVTNSTVLQYMDNLNANDFDNLISLFAEDGALQPPFQKPIVGKENA Syn 6803 MGFTAGKDG KRIAEPV - VPPQEMSQRTKVQIEGVTNSTVLQYMDNLNANDFDLIELFŤSDGALQPPFQRPIVGKENV Nos 7120 MGFDAAKLDGYTRVAEPL - VAPKDISQRVQVTIEGINNSTVLÑYMNNLNANDFDLIKLFVEĎGALQPPFQRPIIGKDAI N. punc 1 MGFDPNAPGSYKKVSEPVAPĎTAPAFRTKVSIEGINNFTVĽGYINNMANDFDAAVALFTŠEGGLQPPFQRPIVGQDAIR Syn 8102 MGYDPVDD - SQVVTEPI - VAPTPVDQREEILIPGVLNQŤŤILSYMQLLNANDFDQLIDLFČNDGALQPPFQRPIVGQDAIR G. viol MGFDPSVVPDEAPEAEDFQFERŤEPVSTERIDVKGVDD PTPLŘÝFEAMNSDNFEAAVALFEPĚGALQPPFQKPIVGREAI N. punc 2 MGFDPSLADKKQAQ - INFKĚPRTS LSPQFTIEGVTEPŤVLKYIEAMNADNFEAAVALFEPĚGALQPPFQKPIVGREAI N. punc 2 MGFDPSLADKKQAQ - INFKĚPRTS LSPQFTIEGVTEPŤVLKYIEAMNADNFEAAVALFANNGALQPPFQKPIVGREAI N. punc 2 MGFDPSLADKKQAQ - INFKĚPRTS - LSPQFTIEGVTEPŤVLKYIEAMNADNFEAAVALFANNGALQPPFQKPIVGREAI N. punc 1 ÅYMREECQULKLIPERGVEPŤEDGYTQIKVTGKVQTPWFGGNVGMNIAWRFLLNPEKVFFVAIDLLASPKELLNL Syn 6803 ĹŘFFREECQNLKLIPERGVEPĚQAEDGTQIKVTGKVQTPWFĞGNVGMNIAWRFLLNPEGKÍŤFVAIDLLASPKELLNLAS Nos 7120 LRĚFFREECQNLKLIPERGVEPĚDAEDGYTQVKVTGKVQTPWFĞGSVGMNIAWRFLLDPEGKÍŤFVAIDLLASPKELLNLVR N. punc 1 ÅYMREECQGLLKMIPERGIŠÉPVEDGYTQVKVTGKVQTPWFĞGSVGMNIAWRFLLDPQGKÍŤFVAIDLLASPKELLNLVR N. punc 1 ÅYMREECQGLKMIPERGIŠÉPVEDGYTQVKVTGKVQTPWFĞGSVGMNIAWRFLLDPQGKÍŤFVAIDLLASPKELLNLVR N. punc 2 TAYĻRDEGQGLVMKPTKGVŠĚTIEJGSKKLKVTGKVQTPWFĞÛNVAMNIAWRFLLDPOKKIYFVAIDLLASPKELLNLVR N. punc 2 TAYĻRDEGQGLVMKPTKGVŠĚTIEJGYTQUKVTGKVQTPWFĞÛNVAMNIAWRFLLDPQGKIFÝVGIDLLASPKELLNLTŘ	G. viol	RVMFDLARRAD	TPISRSYGY	FSVNTKLGFW	YQLAEWMAQG	TVAPIPANYQM	STDAQLLFES	IKNLDGGQQIQ	VLRDIVLN
A. max MGFDTSKLGŞYQRVAEPV - VPPQEMSQRTKVQIEGVTNSTVLQYMDNLNANDFDNLISLFAEDGALQPPFQKPIVGKENA Syn 6803 MGFTAGKDG KRIAEPV - VPPQDTASRTKVSIEGVTNATVLQYMDNLNANDFDTLIELFTSDGALQPPFQRPIVGKENV Nos 7120 MGFDAAKLDGYTRVAEPL - VAPKDISQRVQVTIEGINNSTVLÑYMNNLNANDFDELIKLFVEDGALQPPFQRPIIGKDAI N. punc 1 MGFDPNAPGSYKKVSEPVAPPTAPAFRTKVSIEGINNFTVLQYMUNLNANDFDAAVALFTSSEGGLQPPFQRPIVGQDAIR Syn 8102 MGFDPSVVPDEAPEAEDFQFERTEPVDQREEILIPGVLNQTILSYMQLLNANDFDQLIDLFLNDGALQPPFQRPIVGRDAT G. viol MGFDPSVVPDEAPEAEDFQFERTEPVSTERIDVKGVDDPTPLRYFEAMNSDNFEAAVALFEPEGALQPPFQKPIVGREAT N. punc 2 MGFDPSLADKKQAQ - INFKFPRTS LSPQFTIEGVTEPTVLKYIEAMNADNFEAAVALFANNGALQPPFQKPIVGREAT A. max LRFFREECQNLKLIPERGVSEPTEDGYTQIKVTGKVQTPWFGGNVGMNIAWRFLLNPENKVFFVAIDLLASPKELLNL Syn 6803 N. punc 1 MFFREECQNLKLIPERGVTEPAEDGFTQIKVTGKVQTPWFGGNVGMNIAWRFLLNPEGKTFFVAIDLLASPKELLNL Syn 6803 N. punc 1 MFFREECQNLKLIPERGVAEPADDGYTQVKVTGKVQTPWFGGNVGMNIAWRFLLNPEGKTFFVAIDLLASPKELLNL - Syn 6803 N. punc 1 MFFREECQNLKLIPERGVAEPADDGYTQVKVTGKVQTPWFGGNVGMNIAWRFLLNPEGKTFFVAIDLLASPKELLNLR N. punc 1 MFFREECQNLKLIPERGVAEPADDGYTQVKVTGKVQTPWFGGNVGMNIAWRFLLNPEGKTFFVAIDLLASPKELLNLR N. punc 1 MFFREECQNLKLIPERGVAEPADDGYTQVKVTGKVQTPWFGASVGMNIAWRFLLDPQGKIFFVAIDLLASPKELLNLR N. punc 1 MFFREECQNLKLIPERGVAEPADDGYTQVKVTGKVQTPWFGASVGMNIAWRFLLDPQGKIFFVAIDLLASPKELLNLR N. punc 1 MFFREECQNLKLIPERGVAEPADDGYTQVKVTGKVQTPWFGASVGMNIAWRFLLDPQGKIFFVAIDLLASPKELLNLR N. punc 2 MGFDFREECQNLKLMPERGISEPVEDGYTQVKVTGKVQTPWFGASVGMNIAWRFLLDPQGKIFFVAIDLLASPKELLNLR N. punc 2 MFFREECQNLKLMPQGGYGEPTEGGFNQIKVTGKVQTPWFGASVGMNIAWRFLLDPQGKIFFVAIDLLASPKELLNLR N. punc 2 MFFKRDCQNLKLMPQGGYGEPTEGGFNQIKVTGKVQTPWFGANVAMRFLLDPQGKIFFVAIDLLASPKELLNLR N. punc 2 TAYYRPEEQGLVMKPTKGVSETIEDGYTQKKTGVQTPWFGYNVAMNIAWRFALNPEGKIFYVGIDLLASPKELLNLR N. punc 2 TAYYRPEEQGLVMKPTKGVSETIEDGYTQKKTGVQTPWFGYNVAMNIAWRFALNPEGKIFYVGIDLLASPKELLNLR N. punc 2 TAYYRPEEAQGLTMKPIEGITEV N. punc 2 TAYYRPEAPACPT	N. punc 2	QLMRDLASNAD	TPISRSYAY	FGVNAKLGFW	WQLGEWMKE ¹²⁰	IVAPMPVGYQM	STQVKAVLEA	VQKIDQSQQII	VLRNTVVN
A max MGFDTSKLGŞYQRVAEPV-VPPQEMSQRTKVQIEGVTNSTVLQYMDNLNANDFDNLISLFAEDGALQPFPQRPIVGKENA Syn 6803 MGFTAGKDGKRIAEPV-VPPQDTASRTKVSIEGVTNATÜNNMDNLNANDFDTLIELFÖDGALQPPFQRPIVGKENV Nos 7120 MGFDAAKLDGYTRVAEPL-VAPKDISQRVQVTIEGINNSTVLÖÜYMNNLNANDFDELIKLFVEÖĞALQPPFQRPIIGKDAI N. punc 1 MGFDPNAPGŞYKKVSEPVAPPTAPAFRTKVSIEGINNFTVLÖĞYINNMNANDFDAAVALFTŠEGGLQPPFQRPIVGQDAIR Syn 8102 MGYDPDVDD-SQVVTEPI-VAPFTPVDQREEILIFGVLNQŤILSYMQLLNANDFDQLIDLFLNDGALQPFFQRPIVGQDAIR G. viol MGFDPSVVPDEAPEAEDFQFERTÉPVSTERIDVKGVDDPTPLRYFEAMNŞDNFEAAVALFPÉGALQPFFQKPIVGREAI N. punc 2 MGFDPSLADKKQAQ-INFKFPRTSLSPQFTIEGVTEPŤVLKYIEAMNADNFEAAVALFANNGALQPFFQKPIVGREAI A. max LRFFREECQNLKLIPERGVSEPTEDGYTQIKVTGKVQTPWFGGNVGMNIAWRFLLNPENKVFFVAIDLLASPKELLNL Syn 6803 LRFFREECQNLKLIPERGVEPŽÖDDGTQVKVTGKVQTPWFGGNVGMNIAWRFLLNPEKVFFVAIDLLASPKELLNLVR No punc 1 ÅYMREECQGLLKMIPERGISÉPVEDGYTQVKVTGKVQTPWFGÅAVGMNVAWRFLLDPQGKIFFVAIDLLASPKELLNLVR Syn 8102 LKFFFRDCQNLKLIPERGVEPŽÖDGYTQVKVTGKVQTPWFGANVGMNIAWRFLLNPEGKIFÝVAIDLLASPKELLNLVR No punc 2 TAYLRDEGQLVMKPTKGVŠETIEDGYTQIKVTGKVQTPWFGÅNVGMNIAWRFLLDEDDKVYVIDILASPKELLNLRP Nom 2 TAYLRDEGQGLVMKPTKGVŠETIEDGYTQIKVTGKVQTPWFGÅNVGMNIAWRFLLNPEGKIFÝVGIDLLASPKELLNLRP	250	80		100		120	140		
Syn 6803 MGFTAGKDG KRIAEPV - VPPQDTASRTKVSIEGVTNAT ²⁰ UNYMDNLNANDFDTLIELF ²¹ SDGALQPPFQRPIVGKENV Nos 7120 MGFDAAKLDGYTRVAEPL - VAP ¹ KDISQRVQVTIEGINNSTVL ²⁰ MNNLNANDFDELIKLFVE ²¹ GALQPPFQRPIVGKENV N. punc 1 MGFDPNAPGSYKKVSEPVAP ¹ TAPAFRTKVSIEGINNFTVL ²⁰ UGYINNMNANDFDAAVALFT ²⁵ EGGLQPPFQRPIVGQDAIR Syn 8102 MGYDPUDD - SQVVTEPI - ¹ VAPTPVDQREEILIPGVLNQ ²⁰ ULSYMQLLNANDFDQLIDL ²¹ LNDGALQPPFQRPIVGRDAI G. viol MGFDPSVVPDEAPEAEDFQFER ¹ EPVSTERIDVKGVDDPTPL ²⁰ VFEAMNSDNFEAAVALFEP ²⁶ GALQPPFQKPIVGREAI N. punc 2 MGFDPSLADKKQAQ - INFK ¹ PRTS - LSPQFTIEGVTEP ²¹ VLKYIEAMNADNFEAAVAL ² FANNGALQPPFQKPIVGREAI A. max LRFFREECQNLKLIPERGVSEP ¹ EDGYTQIKVTGKVQTPWFGGNVGMNIAWRFLLNPENKVFFVAIDLLASPKELLNL Syn 6803 ¹ URFFREECQNLKLIPERGV ² P ² DAEDGFTQIKVTGKVQTPW ² GGNVGMNIAWRFLLNPEGK ¹ F ² VAIDLLASPKELLNLAR Nos 7120 LR ² FFREECQNLKLIPERGV ² ²⁰ DGYTQVKVTGKVQTPW ² GGSVGMNIAWRFLLNPQGK ¹ F ² VAIDLLASPKELLNLVR N. punc 1 ²⁴ YMREECQGLLKMIPERGIS ² ²⁰ DFTEGGFNQIKVTGKVQTP ² FGSVGMNIAWRFLLDPQGK ¹ F ² VAIDLLASPKELLNLVR Syn 8102 LKFFKRDCQNLKLMPQGG ² ² CPTEGGFNQIKVTGKVQTP ² FGGEVGMNVAWRFLLDPQGK ¹ F ² VAIDLLASPKELLNLVR N. punc 1 ²⁴ YMREECQGLLKMIPERGIS ² CPTEGGFNQIKVTGKVQTP ² FGGEVGMNVAWRFLLDPQGK ¹ F ² VAIDLLASPKELLNLVR N. punc 1 ²⁴ YMREECQGLLKMIPERGIS ² CPTEGGFNQIKVTGKVQTP ² FGGEVGMNVAWRFLLDPQGK ¹ F ² VAIDLLASPKELLNLVR N. punc 1 ²⁴ YMREECQGLLKMIPERGIS ² CPTEGGFNQIKVTGKVQTP ² FGGEVGMNVAWRFLLDPQGK ¹ F ² VAIDLLASPKELLNLVR N. punc 2 TAYLRDEGQGLVMKPTKG ² C ² FTEGGFNQIKVTGFVETP ² FGGNVGMNIAWRFLLDPQG ² CVYVAIDLLASPKELLNL ² F ² YVGIDLLASPKELLNL ² F ²				100		120	140		
Nos 7120 MGFDAAKLDGYTRVAEPL - VAP ^{ko} DI SQRVQVTI EGINNSTVL ^ŵ YMNNLNANDFDELI KLFVE ²⁰ GALQ PPFQRPI I GKDAI N. punc 1 MGFDPNAPGŞYKKVSEPVAP ^b TAPAFRTKVSIEGINN PTV ²⁰ GYINNMNANDFDAAVALFT ²² SEGGLQ PPFQRPI VGQDAIR Syn 8102 MGYDPVDD - SQVVTEPI - ¹ ^ŵ APTPVDQREEILĮ PGVLNQ ²¹ I LSYMQLLNANDFDQLIDL ²² LNDGALQ PPFQRPI VGRDA ¹ ^c G. viol ^M GFDPSVVPDEAPEAEDFQFERT ¹ ^e PVSTERIDVKGVDD PTPL ² ^ŵ YFGGNVGMNIAWRFLDGALQ PPFQKPI VGREAI N. punc 2 MGFDPSLADKKQAQ - INFK ¹ ^b PRTS LSPQFTI EGVTEP ¹ ^w VKYIEAMNADNFEAAVALF ²² ⁰ ANNGALQ PPFQKPI VGREAI ¹⁰⁰ A. max LRFFREECQNLKLI PERGVSEPTEDGYTQI KVTGKVQ TPWFGGNVGMNIAWRFLLNPENKVFFVAIDLLASPKELLNL Syn 6803 LR ² ^h FFREECQNLKLI PERGVT ² ^b PAEDGFTQI KVTGKVQ TPW ² ^b GGNVGMNIAWRFLLNPEGKI ¹ ^b FVAIDLLASPKELLNLVR Nos 7120 LR ² ^h FFREECQNLNLL PERGVAE ² ^b ADDGYTQVKVTGKVQ TPW ² ^b GASVGMNIAWRFLLNPQGKI ¹ ^b FVAIDLLASPKELLNLVR N. punc 1 ²⁴⁰ ^A YMREECQGLLKMI PERGIS ² ^b PVEDGYTQVKVTGKVQ TPW ² ^b GASVGMNIAWRFLLDPQGK ¹ ¹ ^b VAIDLLASPKELLNLVR Syn 8102 LKFFKRDCQNLKLMPQGY ² ^G BPTEGGFNQI KVTGKVQ TPW ² ^b GASVGMNVAWRFL LDEND ³ ^k VYAIDLLASPKELLNLVR N. punc 2 TAYLRDEGQGLVMKPTKG ² ⁸ ^b ETIEDGYTQHKITGTVETP ³ ^k FGGNVGMNIAWRFLLNPQG ¹ ¹ ^b VAIDLLASPKELLNL ² ¹ ^b	A. max	MGFDTSKLGSY	QRVAEPV-V	PPQEMSQRTK	VQIEGVTNST	VLQYMDNLNAN	DFDNLISLFA	EDGALQPPFQK	PIVGKENA
N. punc 1 MGFDPNAPGŞYKKVSEPVAP ³⁸⁹ TAPAFRTKVSIEGINNFTV ²⁰⁰ GYINNMNANDFDAAVALFT ²⁵⁰ EGGLQPPFERPIVGQDAIR Syn 8102 MGYDPDVDD - SQVVTEPI - ¹⁶⁰ VAPTPVDQREEILĮPGVLNQ ²⁰¹ LSYMQLLNANDFDQLIDLF ²⁵⁰ LNDGALQPPFQRPIVGRDA ²⁴⁰ G, viol MGFDPSVVPDEAPEAEDFQFERTEPVSTERIDVKGVDDPTPL ²⁶⁰ VFEAMNSDNFEAAVALFEP ²⁵⁰ GALQPPFQKPIVGREAI N. punc 2 MGFDPSLADKKQAQ - INFK ⁵⁶ PRTS - LSPQFTIEGVTEP ²⁷⁰ VLKYIEAMNADNFEAAVALF ²⁶⁰ ANNGALQPPFQKPIVGREAI A. max LRFFREECQNĻKLIPERGVSEPTEDGYTQIKVTGKVQTPWFGGNVGMNIAWRFLLNPENKVFFVAIDLLASPKELLNL Syn 6803 LRFFREECQNĻKLIPERGVT ²⁶⁰ PAEDGFTQIKVTGKVQTPWFGGNVGMNIAWRFLLNPEGK ³⁶⁰ FFVAIDLLASPKELLNLAR Nos 7120 LR ²⁴⁰ FREECQNĻNLLPERGVAEP ²⁶⁰ DGYTQVKVTGKVQTPWFGASVGMNIAWRFLLNPQGKI ⁵⁶⁷ VAIDLLASPKELLNLVR N. punc 1 ÅYMREECQGĻLKMIPERGIS ²⁶⁰ PVEDGYTQVKVTGKVQTPW ²⁶⁰ GASVGMNIAWRFLLDPQGK ³⁶⁰ FFVAIDLLASPKELLNLVR Syn 8102 LKFFKRDCQNLKĻMPQGGY ²⁶⁰ EPTEGGFNQIKVTGKVQTP ²⁶⁰ GREVGMNVAWRFLLDPQGK ³⁶⁰ FFVAIDLLASPKELLNLVR N. punc 2 TAYĻRDEGQGLVMĶPTKGV ²⁶⁰ ETIEDGYTQHKITGTVETP ²⁶⁰ FGGNVGMNIAWRFLLNPEGKI ⁵⁷⁰ VGIDLLASPKELLNLR ³⁵⁰ N. punc 2 TAYĻRDEGQGLVMĶPTKGV ²⁶⁰ ETIEDGYTQHKITGTVETP ²⁶⁰ FGGNVGMNIAWRFLNPG ²⁶⁰ YFVAIDLLASPKELLNLT ⁴⁷	A. max Syn 6803	MGFDTSKLGSY MGFTAGKDG	QRVAEPV-V ĶRIAEPV-V	PPQEMSQRTK 190 PPQDTASRTK	VQIEGVTNST VSIEGVTNAT	VLQYMDNLNAŅ VLQYMDNLNAŅ VLŅYMDNLNAN	DFDNLISLFA DFDTLIELFT	EDGALQPPFQK SDGALQPPFQK	CPIVGKENA PIVGKENV
Syn 8102 MGYDPDVDD - SQVVTEPI - ¹ ¹ ⁰ ⁰ APTPVDQREEILĮPGVLNQ ²¹ ILSYMQLLNANDFDQLIDL ²² ⁰ LNDGALQPPFQRPIVGRDA ¹ G. viol ¹ ¹⁰ ⁰ MGFDPSVVPDEAPEAEDFQFER ¹ ¹⁰ ⁰ PPVSTERIDVKGVDDPTPL ²⁰ ¹⁰ YFEAMNSDNFEAAVALFEP ² ¹⁰ GALQPPFQKPIVGREAI N. punc 2 MGFDPSLADKKQAQ - INFK ¹ ¹⁰ PRTS LSPQFTIEGVTEP ²¹ ¹⁰ VLKYIEAMNADNFEAAVAL ² ¹⁰ ¹⁰ A. max LRFFREECQNLKLIPERGVSEPTEDGYTQIKVTGKVQTPWFGGNVGMNIAWRFLLNPENKVFFVAIDLLASPKELLNL Syn 6803 ¹¹ ¹⁰ ¹⁰ Nos 7120 LR ²¹ ¹⁰ FREECQNLKLIPERGVAEP ²⁰ ²⁰ DDGYTQVKVTGKVQTPWF ² GAVGMNMAWRFLLNPEGK ¹⁰ ¹⁰ FVAIDLLASPKELLNLVR N. punc 1 ²⁴⁰ AYMREECQGLLKMIPERGIS ²⁰ PVEDGYTQVKVTGKVQTPW ²⁰ GASVGMNIAWRFLLDPQGK ¹⁰ FFVAIDLLASPKELLNLVR Syn 8102 LKFFKRDCQNLKLMPQGG ² ²⁰ GEPTEGGFNQIKVTGKVQTP ²⁰ ²⁰ GREVGMNVAWRFLLDPQGK ¹⁰ FFVAIDLLASPKELLNLVR N. punc 2 ²⁴⁰ MREEAQGLTMKPIEGITEV ²¹ ²⁰ PDGSKKLKVTGKVQTPWF ²⁰ MVAMNIAWRFALNPEGKI ²⁰ ²⁰ VGIDLLASPQELLNLR ³ ³⁰ N. punc 2 TAYLRDEGQGLVMKPTKG ²⁰ ²⁰	A. max Syn 6803 Nos 7120	MGFDTSKLGSY MGFTAGKDG MGFDAAKLDGY	QRVAEPV-V ĶRIAEPV-V TRVAEPL-V	PPQEMSQRTK ¹⁸⁰ PPQDTASRTK APKDISQRVQ	VQIEGVTNST VSIEGVTNAT VTIEGINNST	vlqymdnlnan vlqymdnlnan vlnymdnlnan vlnymnnlnan	DFDNLISLFA DFDTLIELFT DFDTLIELFT DFDELIKLFV	EDGALQPPFQK SDGALQPPFQK EDGALQPPFQK	CPIVGKENA PIVGKENV PIIGKDAI
G. viol MGFDPSVVPDEAPEAEDFQFERITEPVSTERIDVKGVDDPTPLRÖYFEAMNSDNFEAAVALFEPZÖGALQPPFQKPIVGREAI N. punc 2 MGFDPSLADKKQAQ - INFKFPRTS - LSPQFTIEGVTEPTVLKYIEAMNADNFEAAVALFANNGALQPPFQKPIVGREAI A. max LRFFREECQNLKLIPERGVSEPTEDGYTQIKVTGKVQTPWFGGNVGMNIAWRFLLNPENKVFFVAIDLLASPKELLNL Syn 6803 LRFFREECQNLKLIPERGVTEPAEDGFTQIKVTGKVQTPWFGGNVGMNIAWRFLLNPEGKI Nos 7120 LRFFREECQNLNLLPERGVAEPÄDDGYTQVKVTGKVQTPWFGÅAVGMNMAWRFLLNPGKIF Nos 7120 LRFFREECQNLKLIPERGVAEPÄDDGYTQVKVTGKVQTPWFGÅAVGMNMAWRFLLNPQGKIFFVAIDLLASPKELLNLVR N. punc 1 ÅYMREECQGLLKMIPERGISË PVEDGYTQVKVTGKVQTPWFGÅAVGMNVAWRFLLDPQGKIFFVAIDLLASPKELLNLVR Syn 8102 LKFFKRDCQNLKLMPQGGYGEPTEGGFNQIKVTGKVQTPWFGÅAVGMNVAWRFLLDPQGKIFFVAIDLLASPKELLNLVR N. punc 2 TAYLRDEGQGLVMKPTKGVŠETIEDGYTQHKITGTVETPÅFGGNVGMNIAWRFLLNPEGKIFÝVGIDLLASPQELLNLR	A. max Syn 6803 Nos 7120 N. punc 1	MGFDTSKLGSY MGFTAGKDG MGFDAAKLDGY MGFDPNAPGSY	QRVAEPV - V ĶRIAEPV - V TRVAEPL - V KKVSEPVAP	PPQEMSQRTK PPQDTASRTK APKDISQRVQ i ³⁶⁰ PTAPAFRTKY	VQIEGVTNST VSIEGVTNAT VTIEGINNST SIEGINNPTY	VLQYMDNLNAŅ VLQYMDNLNAŅ VLŅYMDNLNAŅ 200 LGYINNMNAŅD	DFDNLISLFA DFDTLIELFT DFDELIKLFV FDAAVALFTS	EDGALQPPFQR SDGALQPPFQR EDGALQPPFQR EGGLQPPFERF	XPIVGKENA RPIVGKENV RPIIGKDAI PIVGQDAIR
N. punc 2 MGFDPSLADKKQAQ - INFK ¹⁶⁰ PRTS - LSPQFTIEGVTEP ²⁰⁰ VLKYIEAMNADNFEAAVAL ²⁷⁰ ANNGALQPPFQKPIVGREA ²¹⁰ A. max LRFFREECQNLKLIPERGVSEPTEDGYTQIKVTGKVQTPWFGGNVGMNIAWRFLLNPENKVFFVAIDLLASPKELLNL Syn 6803 LRFFREECQNLKLIPERGVTE ²⁴⁰ DAEDGFTQIKVTGKVQTPW ²⁶⁰ GGNVGMNIAWRFLLNPEGK ²⁰⁰ FFVAIDLLASPKELLNFAR Nos 7120 LR ²⁴⁰ FREECQNLKLIPERGVAEP ²⁴⁰ DGYTQVKVTGKVQTPW ²⁶⁰ GAVGMNMAWRFLLNPQGKIF ²⁷⁰ VAIDLLASPKELLNLVR N. punc 1 Å ²⁴⁰ MREECQGLLKMIPERGIS ²⁶⁰ PTEGGFNQIKVTGKVQTP ²⁶⁰ GASVGMNIAWRFLLDPQGK ²¹⁰ FFVAIDLLASPKELLNLVR Syn 8102 LKFFKRDCQNLKLMPQGGY ²⁶⁰ EPTEGGFNQIKVTGKVQTP ²⁶⁰ GREVGMNVAWRFLLDEND ⁶⁰ YFVAIDLLASPKELLNLVR G. viol AA ²⁴⁰ MREEAQGLTMKPIEGITEV ²⁶⁰ PDGSKKLKVTGKVQTPWFG ²⁰ NVAMNIAWRFALNPEGKIF ²⁷⁰ VGIDLLASPQELLNLRP N. punc 2 TAYLRDEGQGLVMKPTKGV ²⁵⁰ ETIEDGYTQHKITGTVETP ²⁶⁰ FGGNVGMNIAWRFLDPQ ²⁰ YFVAIDLLASPKELLNLT ²⁷⁰	A. max Syn 6803 Nos 7120 N. punc 1 Syn 8102	™GFDTSKLGSY MGFTAGKDG MGFDAAKLDGY MGFDPNAPGSY MGYDPDVDD-S	QRVAEPV - V KRIAEPV - V TRVAEPL - V KKVSEPVAP QVVTEPI - V	PPQEMSQRTK PPQDTASRTK APKDISQRVQ PTAPAFRTKY APTPVDQREE	VQIEGVTNST VSIEGVTNAT VTIEGINNST SIEGINNPTV ILIPGVLNQT	VLQYMDNLNAŅ ŴLNYMDNLNAŅ ĽĜYINNMLNAŅ LGYINNMNAŅD ILŞYMQLLNAN	DFDNLISLFA DFDTLIELFT DFDELIKLFV FDAAVALFTS DFDQLIDLFL	EDGALQPPFQF SDGALQPPFQF EDGALQPPFQF EGGLQPPFERF NDGALQPPFQF	CPIVGKENA PIVGKENV PIIGKDAI PIVGQDAIR PIVGRDAI
A. max LRFFREECQN LKLIPERGVSEPTEDGYTQI KVTGKVQTPWFGGNVGMNIAWRFLLNPENKVFFVAIDLLASPKELLNL Syn 6803 LRFFREECQN LKLIPERGVTÉPAEDGFTQI KVTGKVQTPWFGGNVGMNIAWRFLLNPEGKIFFVAIDLLASPKELLNFAR Nos 7120 LRFFREECQN LNLLPERGVAEPADDGYTQVKVTGKVQTPWFGÅAVGMNMAWRFLLNPQGKIFFVAIDLLASPKELLNLVR N. punc 1 ÅYMREECQGL LKMIPERGISÉPVEDGYTQVKVTGKVQTPWFGÅAVGMNIAWRFLLDPQGKIFFVAIDLLASPKELLNLVR Syn 8102 LKFFKRDCQN LKLMPQGGYÅEPTEGGFNQI KVTGKVQTPWFGÅNVAMNIAWRFLLDENDÅ G. viol AAÝMREEAQGLTMKPIEGITEVLPDGSKKLKVTGKVQTPWFG N. punc 2 TAYLRDEGQGLVMKPTKGVŠETIEDGYTQHKITGTVETPÅ WFGGNVGMNIAWRFLNPQGQ 1 1 1 1 1 1 1 1 1 1 1 1 1	A. max Syn 6803 Nos 7120 N. punc 1 Syn 8102 G. viol	MGFDTSKLGSY MGFTAGKDG MGFDAAKLDGY MGFDPNAPGSY MGFDPDVDD - S	QRVAEPV - V KRIAEPV - V TRVAEPL - V KKVSEPVAP QVVTEPI - V APEAEDFQF	PPQEMSQRTK PPQDTASRTK APKDISQRVQ PTAPAFRTKV APTPVDQREE ERTEPVSTER	VQIEGVTNST VSIEGVTNAT VTIEGINNST SIEGINNPT ILIPGVLNQT IDVKGVDDPT	VLQYMDNLNAŅ VLQYMDNLNAŅ LÖYMNNLNAŅ LĞYINNMNAŅD ILSYMQLLNAN PLŔYFEAMNŞD	DFDNLISLFA DFDTLIELFT DFDELIKLFV FDAAVALFTS DFDQLIDLFL NFEAAVALFE	EDGALQPPFQR SDGALQPPFQR EDGALQPPFQR EGGLQPPFERE NDGALQPPFQR PEGALQPPFQR	PIVGKENA PIVGKENV PIIGKDAI PIVGQDAIR PIVGRDAI CPIVGREAI
A max LRFFREECQN LKLIPERGVSEPTEDGYTQI KVTGKVQ TPWFGGNVGMN IAWRFLLNPENKVFFVAIDLLASPKELLNL Syn 6803 LRFFREECQN LKLIPERGVTÉPAEDGFTQI KVTGKVQ TPWFGGNVGMN IAWRFLLNPEGKIPFVAIDLLASPKELLNFAR Nos 7120 LRFFREECQN LNLLPERGVAE PÅDDGYTQV KVTGKVQ TPWFGÅAVGMNMAWRFLLNPQGKIPFVAIDLLASPKELLNLVR N. punc 1 ÅYMREECQGL LKMIPERGISÉPVEDGYTQV KVTGKVQ TPWFGÅAVGMN IAWRFLLDPQGKIPFVAIDLLASPKELLNLVR Syn 8102 LKFFKRDCQN LKLMPQGGYGEPTEGGFNQI KVTGKVQ TPWFGÅAVGMNVAWRFLLDENDKI YFVAIDLLASPKELLNLVR G. viol AAYMREEAQGLTMKPIEGITEVLPDGSKKLKVTGKVQTPWFGVNVAMN IAWRFALNPEGKIPVOGIDLLASPQELLNLP N. punc 2 TAYLRDEGQGLVMKPTKGVSETIEDGYTQHKITGTVETPWFGVOMNIAWRFLLNPQGY	A. max Syn 6803 Nos 7120 N. punc 1 Syn 8102 G. viol N. punc 2	MGFDTSKLGSY MGFTAGKDG MGFDAAKLDGY MGFDPNAPGSY MGYDPDVDD-S MGFDPSVVPDE MGFDPSLADKK	QRVAEPV - V KRIAEPV - V TRVAEPL - V KKVSEPVAP QVVTEPI - V APEAEDFQ QAQ - INFKF	PPQEMSQRTK PPQDTASRTK APKDISQRVQ PTAPAFRTKV APTPVDQREE ERTEPVSTER PRTSLSPQ	VQIEGVTNST VSIEGVTNAT VTIEGINNST SIEGINNPTV ILIPGVLNQT IDVKGVDDPT FTIEGVTEPT	VLQYMDNLNAN VLQYMDNLNAN VLÑYMNNLNAN LGYINNMNAND ILSYMQLLNAN PLRYFEAMNSD VLKYIEAMNAD	DFDNLISLFA DFDTLIELFT DFDELIKLFV FDAAVALFTS DFDQLIDLFL NFEAAVALFE NFEAAVALFA	EDGALQPPFQK SDGALQPPFQF EDGALQPPFQF EGGLQPPFFFF NDGALQPPFQF PEGALQPPFQK	EPIVGKENA PIVGKENV PIIGKDAI PIVGQDAIR PIVGRDAI EPIVGREAI
Syn 6803 LRFFREECQNLKLIPERGVTÉPAEDGFTQIKVTGKVQTPWFÖGGNVGMNIAWRFLLNPEGKIFFVAIDLLASPKELLNFAR Nos 7120 LRFFREECQNLNLLPERGVAEPADDGYTQVKVTGKVQTPWFGÅAVGMNMAWRFLLNPQGKIFFVAIDLLASPKELLNLVR N. punc 1 ÅYMREECQGLLKMIPERGISÉPVEDGYTQVKVTGKVQTPWFGÅAVGMNIAWRFLLDPQGKIFFVAIDLLASPKELLNLVR Syn 8102 LKFFKRDCQNLKLMPQGGYÖEPTEGGFNQIKVTGKVQTPWFGVNVAWRFLLDENDKUYFVAIDLLASPKELLKLGÖ G. viol AAYMREEAQGLTMKPIEGITEVLPDGSKKLKVTGKVQTPWFGVNVAMNIAWRFALNPEGKIFVVGIDLLASPCELLNLRP N. punc 2 TAYLRDEGQGLVMKPTKGVSETIEDGYTQHKITGTVETPWFGVGNVGMNIAWRFLLNPQGQUYFVAIDLLASPKELLNLTÅ	A. max Syn 6803 Nos 7120 N. punc 1 Syn 8102 G. viol N. punc 2	MGFDTSKLGSY MGFTAGKDG MGFDAAKLDGY MGFDPNAPGSY MGYDPDVDD-S MGFDPSVVPDE MGFDPSLADKK	QRVAEPV - V KRIAEPV - V TRVAEPL - V KKVSEPVAP QVVTEPI - V APEAEDFQF QAQ - INFKF	PPQEMSQRTK PPQDTASRTK APKDISQRVQ PTAPAFRTKV APTPVDQREE ERTEPVSTER PRTS LSPQ	VQIEGVTNST VSIEGVTNAT VTIEGINNST SIEGINNPT ILIPGVLNQT IDVKGVDDPT FTIEGVTEPT	VLQYMDNLNAN VLQYMDNLNAN VLÑYMNNLNAN LGYINNMNAND ILSYMQLLNAN PLRYFEAMNSD VLKYIEAMNAD	DFDNLISLFA DFDTLIELFT DFDELIKLFV FDAAVALFTS DFDQLIDLFL NFEAAVALFE NFEAAVALFE	EDGALQPPFQR SDGALQPPFQR EDGALQPPFQR EGGLQPPFERF NDGALQPPFQR PÊGALQPPFQR NNGALQPPFQR 220	PIVGKENA PIVGKENV PIIGKDAI PIVGQDAIR PIVGRDAI PIVGREAI CPIVGREAI
Nos 7120 LRFFREECQN LNLL PERGVAE PÅDDGYTQV KVTGKVQT PWFGÅAVGMNMAWRFLLN PQGKI FFVAIDLLAS PKELLNLVR N. punc 1 ÅYMREECQGLLKMI PERGI SËPVEDGYTQV KVTGKVQT PWFGÅAVGMN I AWRFLLD PQGKI FFVAIDLLAS PKELLNLVR Syn 8102 LKFFKRDCQN LKLM PQGGYGE PTEGGFNQI KVTGKVQT PWFGÅVGMN VAWRFLLDENDŘI YFVAIDLLAS PKELLNLVR G. viol AAÝMREEAQGLTMKPIEGITEVĽ PDGSKKLKVTGKVQT PWFGVNVAMN I AWRFALN PEGKI FÝVGIDLLAS PQELLNLRP N. punc 2 TAYLRDEGQGLVMĶPTKGVŠETIEDGYTQHKITGTVET PWFGVNVAMN I AWRFLN PQGQ YVAIDLLAS PKELLNLTŘ	A. max Syn 6803 Nos 7120 N. punc 1 Syn 8102 G. viol N. punc 2 A. max	MGFDTSKLGSY MGFTAGKDG MGFDAAKLDGY MGFDPNAPGSY MGYDPDVDD-S MGFDPSVVPDE MGFDPSLADKK	QRVAEPV - V KRIAEPV - V TRVAEPL - V KKVSEPVAP QVVTEPI - V APEAEDFQF QAQ - INFKF	PPQEMSQRTK PPQDTASRTK APKDISQRVQ PTAPAFRTKV APTPVDQREE ERTEPVSTER PRTS LSPQ EPTEDGYTQI	VQIEGVTNST VSIEGVTNAT VTIEGINNST SIEGINNPTV ILIPGVLNQT IDVKGVDDPT FTIEGVTEPT	VLQYMDNLNAN VLQYMDNLNAN VLÑYMNNLNAŅ LGYINNMNAŅD ILSYMQLLNAN PLRYFEAMNŞD VLKYIEAMNAD 200 FGGNVGMNIAW	DFDNLISLFA DFDTLIELFT DFDELIKLFV FDAAVALFTS DFDQLIDLFL NFEAAVALFE NFEAAVALFA NFEAAVALFA	EDGALQPPFQR SDGALQPPFQR EDGALQPPFQR EGGLQPPFQR PEGALQPPFQR PEGALQPPFQR NNGALQPPFQR 220	PIVGKENA PIIGKDAI PIVGQDAIR PIVGQDAIR PIVGREAI PIVGREAI
N. punc 1 ÅYMREECQGLLKMIPERGISÉPVEDGYTQVKVTGKVQTPWFGASVGMNIAWRFLLDPQGKŮFFVAIDLLASPKELLNLVR Syn 8102 LKFFKRDCQNLKLMPQGGYĠEPTEGGFNQIKVTGKVQTPŴFGREVGMNVAWRFLLDENDŘIYFVAIDLLASPAELLKLGĠ G. viol AAŸMREEAQGLTMKPIEGITEVĹPDGSKKLKVTGKVQTPWFGŮNVAMNIAWRFALNPEGKIFŸVGIDLLASPQELLNLRP N. punc 2 TAYLRDEGQGLVMĶPTKGVŠETIEDGYTQHKITGTVETPŴFGGNVGMNIAWRFLLNPQGŮYFVAIDLLASPKELLNLTŘ	A. max Syn 6803 Nos 7120 N. punc 1 Syn 8102 G. viol N. punc 2 A. max Syn 6803	MGFDTSKLGSY MGFDTAGKDG MGFDAAKLDGY MGFDPNAPGSY MGYDPDVDD-S MGFDPSVVPDE MGFDPSLADKK 100 LRFFREECQNL LRFFREECQNL	QRVAEPV - V KRIAEPV - V TRVAEPL - V KKVSEPVAP QVVTEPI - V APEAEDFQ QQQ - INFKF KLIPERGVS KLIPERGVT	PPQEMSQRTK PPQDTASRTK APKDISQRVQ PTAPAFRTKV APTPVDQREE ERTEPVSTER PRTS - LSPQ EPTEDGYTQI EPAEDGFTQI	VQIEGVTNST VSIEGVTNAT VTIEGINNST SIEGINNPV ILIPGVLNQT IDVKGVDDPT FTIEGVTEPT KVTGKVQTPW	VLQYMDNLNAN VLQYMDNLNAN VLÑYMNNLNAN LĜYINNMNAND ILSYMQLLNAN PLŘYFEAMNSD VLKYIEAMNAD FGGNVGMNIAW	DFDNLISLFA DFDTLIELFT DFDELIKLFV FDAAVALFTS DFDQLIDLFL NFEAAVALFE NFEAAVALFA RFLLNPENKV RFLLNPEGKI	EDGALQPPFQK SDGALQPPFQK EDGALQPPFQK EGGLQPPFQK PEGALQPPFQK PEGALQPPFQK 220 FFVAIDLLASE	PIVGKENA PIVGKENV PIIGKDAI PIVGQDAIR PIVGRDAI PIVGREAI PIVGREAI PKELLNL PKELLNFAR
Syn 8102 LKFFKRDCQNLKLMPQGGY ² ⁶ ⁶ EPTEGGFNQIKVTGKVQTP ² ⁶ ⁶ FGREVGMNVAWRFLLDEND ² ⁶ NYAIDLLASPAELLKLG ³ ⁶ G. viol AA ² ⁶ MREEAQGLTMKPIEGITEV ² ⁶ PDGSKKLKVTGKVQTPWFG ² ⁶ NVAMNIAWRFALNPEGKIF ² ⁶ VGIDLLASPQELLNLRP N. punc 2 TAYLRDEGQGLVMKPTKGV ² ⁶ ETIEDGYTQHKITGTVETP ² ⁶ FGGNVGMNIAWRFLLNPQG ⁰ ₀ IYFVAIDLLASPKELLNLT ²	A. max Syn 6803 Nos 7120 N. punc 1 Syn 8102 G. viol N. punc 2 A. max Syn 6803 Nos 7120	MGFDTSKLGSY MGFDTAGKDG MGFDAAKLDGY MGFDPNAPGSY MGFDPSVVPDE MGFDPSLADKK MGFDPSLADKK LRFFREECQNL LRFFREECQNL LRFFREECQNL LRFFREECQNL	QRVAEPV - V KRIAEPV - V TRVAEPL - V KKVSEPVAP QVVTEPI - V APEAEDFQF QAQ - INFKF KLIPERGVS KLIPERGVT NLLPERGVA	PPQEMSQRTK PPQDTASRTK APKDISQRVQ PTAPAFRTKV APTPVDQREE ERTEPVSTER PRTS - LSPQ EPTEDGYTQI EPAEDGFTQI EPAEDGFTQI	VQIEGVTNST VSIEGVTNAT VTIEGINNST SIEGINNPTV ILIPGVLNQT IDVKGVDDPT FTIEGVTEPT KVTGKVQTPW KVTGKVQTPW	VLQYMDNLNAŅ VLQYMDNLNAŅ LGYINNMNAŅD ILSYMQLLNAN PLŔŶFEAMNŞD VLKYIEAMNAD FGGNVGMNIAW FGÂAVGMNMAW	DFDNLISLFA DFDTLIELFT DFDELIKLFV FDAAVALFTS DFDQLIDLF NFEAAVALFE NFEAAVALFE NFEAAVALFA RFLLNPENKV RFLLNPEGKI RFLLNPGKI	EDGALQPPFQF SDGALQPPFQF EDGALQPPFQF EGGLQPPFQF PEGALQPPFQF NDGALQPPFQF NNGALQPPFQF FFVAIDLLASE FFVAIDLLASE	PIVGKENA PIVGKENV PIIGKDAI PIVGQDAIR PIVGRDAI CPIVGREAI PEVGREAI
G. viol AAYMREEAQGLTMKPIEGITEVLPDGSKKLKVTGKVQTPWFGVNVAMNIAWRFALNPEGKIFYVGIDLLASPQELLNLRP N. punc 2 TAYLRDEGQGLVMKPTKGVSETIEDGYTQHKITGTVETPWFGGNVGMNIAWRFLLNPQGQ1YFVAIDLLASPKELLNLTR	A. max Syn 6803 Nos 7120 N. punc 1 Syn 8102 G. viol N. punc 2 A. max Syn 6803 Nos 7120 N. punc 1	MGFDTSKLGSY MGFDTSKLGSY MGFDAAKLDGY MGFDPNAPGSY MGFDPNAPGSY MGFDPSVVPDE MGFDPSLADKK MGFDPSLADKK LRFFREECQNL LRFFREECQNL LRFFREECQNL AYMREECQGLL	QRVAEPV - V KRIAEPV - V TRVAEPL - V KKVSEPVAP QVVTEPI - V APEAEDFQF QAQ - INFKF KLIPERGVS KLIPERGVT NLLPERGVA KMIPERGIS	PPQEMSQRTK PPQDTASRTK APKDISQRVQ PTAPAFRTKV APTPVDQREE ERTEPVSTER PRTS LSPQ EPTEDGYTQI EPAEDGFTQI EPAEDGYTQV EPVEDGYTOV	VQIEGVTNST VSIEGVTNAT VTIEGINNST SIEGINNPTV ILIPGVLNQT IDVKGVDDPT FTIEGVTEPT KVTGKVQTPW KVTGKVQTPW KVTGKVQTPW	VLQYMDNLNAN VLQYMDNLNAN UZYMNNLNAN LQYINNMNAND ILSYMQLLNAN PLRYFEAMNSD VLKYIEAMNAD VLKYIEAMNAD FGGNVGMNIAW FGAAVGMNNAW	DFDNLISLFA DFDTLIELFT DFDELIKLFV FDAAVALFTS DFDQLIDLFL NFEAAVALFE NFEAAVALFA RFLLNPENKV RFLLNPEGKI RFLLNPQGKI	EDGALQPPFQF SDGALQPPFQF EDGALQPPFQF EGGLQPPFQF PEGALQPPFQF PEGALQPPFQF NNGALQPPFQF SFFVAIDLLASF FFVAIDLLASF FFVAIDLLASF	PIVGKENA PIVGKENV PIIGKDAI PIVGQDAI PIVGRDAI CPIVGREAI CPIVGREAI PIVGREAI PKELLNL PKELLNLAP PKELLNLVR
N. punc 2 TAYLRDEGQGLVMKPTKGVSETIEDGYTQHKITGTVETPWFGGNVGMNIAWRFLLNPQGQIYFVAIDLLASPKELLNLTR	A. max Syn 6803 Nos 7120 N. punc 1 Syn 8102 G. viol N. punc 2 A. max Syn 6803 Nos 7120 N. punc 1 Syn 8102	MGFDTSKLGSY MGFDTSKLGSY MGFDAAKLDGY MGFDPNAPGSY MGFDPSVVPDE MGFDPSVVPDE MGFDPSLADKK ito LRFFREECQNL 240 LRFFREECQNL 240 LRFFREECQNL 240 AYMREECQGLL LKFFKRDCONL	QRVAEPV - V KRIAEPV - V TRVAEPL - V KKVSEPVAP QVVTEPI - ^V APEAEDFQF QAQ - INFKF KLIPERGVS KLIPERGVT NLLPERGVA KMIPERGIS KLMPQGGY ²⁶	PPQEMSQRTK PPQDTASRTK APKDISQRVQ PTAPAFRTKV APTPVDQREE ERTEPVSTER PRTS LSPQ EPTEDGYTQI EPAEDGFTQI EPAEDGFTQI EPAEDGYTQV EPVEDGYTQV EPTEGGFNOI	VQIEGVTNST VSIEGVTNAT VTIEGINNST SIEGINNPTV ILIPGVLNQT IDVKGVDDPT FTIEGVTEPT KVTGKVQTPW KVTGKVQTPW KVTGKVQTPW KVTGKVQTPW	VLQYMDNLNAN VLQYMDNLNAN LOYMNNLNAN LOYINNMNAND ILSYMQLLNAN PLRYFEAMNSD VLKYIEAMNAD FGGNVGMNIAW FGGNVGMNIAW FGASVGMNIAW FGASVGMNIAW	DFDNLISLFA DFDTLIELFT DFDELIKLFV FDAAVALFTS DFDQLIDL 220 DFDQLIDL 220 NFEAAVALFE NFEAAVALFA RFLLNPENKV RFLLNPEKT RFLLNPQGKI RFLLDPQGKI RFLLDPQGKI	EDGALQPPFQF SDGALQPPFQF EDGALQPPFQF EGGLQPPFQF PEGALQPPFQF PEGALQPPFQF SFVAIDLLASF FFVAIDLLASF FFVAIDLLASF FFVAIDLLASF	PIVGKENA PIVGKENV PIIGKDAI PIVGQDAIR PIVGRDAI CPIVGREAI CPIVGREAI CPIVGREAI CPIVGREAI CPIVGREAI
	A. max Syn 6803 Nos 7120 N. punc 1 Syn 8102 G. viol N. punc 2 A. max Syn 6803 Nos 7120 N. punc 1 Syn 8102 G. viol	MGFDTSKLGSY MGFDTSKLGSY MGFDAAKLDGY MGFDPNAPGSY MGFDPSVVPDE MGFDPSVVPDE MGFDPSLADKK LRFFREECQNL LRFFREECQNL LRFFREECQNL LRFFREECQNL LRFFREECQNL LKFFREECQGLL LKFFKRDCQNL AAYMREECQGL	QRVAEPV - V KRIAEPV - V TRVAEPL - V KKVSEPVAP QVVTEPI - V APEAEDFQF QAQ - INFKF KLIPERGVS KLIPERGVS KLIPERGVA KMIPERGIS KLMPQGGY2 TMKPIEGIT	PPQEMSQRTK PPQDTASRTK APKDISQRVQ PTAPAFRTKV APTPVDQREE ERTEPVSTER PRTS LSPQ EPTEDGYTQI EPADGYTQV EPVEDGYTQV EPVEDGYTQV EPTEGGFNQI EVLPDGSKKL	VQIEGVTNST VSIEGVTNAT VTIEGINNST SIEGINNPTV ILIPGVLNQT IDVKGVDDPT FTIEGVTEPT KVTGKVQTPW KVTGKVQTPW KVTGKVQTPW KVTGKVQTPW	VLQYMDNLNAN VLQYMDNLNAN VLÑYMNNLNAŅ LĜYINNMNAŅD ILSYMQLLNAN PLRŶFEAMNSD VLKYIEAMNAD FGGNVGMNIAW FGÂAVGMNIAW FGÂSVGMNIAW FGREVGMNVAW FGŶNVAMNIAW	DFDNLISLFA DFDTLIELFT DFDELIKLFV FDAAVALFTS DFDQLIDLFL NFEAAVALFE NFEAAVALFA RFLLNPENKV RFLLNPEGKI RFLLNPQGKI RFLLDPQGKI RFLLDPQGKI	EDGALQPPFQR SDGALQPPFQR EDGALQPPFQR EGGLQPPFQR PEGALQPPFQR PEGALQPPFQR S20 FFVAIDLLASE FFVAIDLLASE FFVAIDLLASE FFVAIDLLASE FFVAIDLLASE	PIVGKENA PIVGKENV PIIGKDAI PIVGQDAI PIVGRDAI CPIVGREAI CPIVGREAI CPIVGREAI CPIVGREAI

Abbreviations: A. max, Anthrospira maxima [11]; Syn 6803, Synechocystis PCC 6803 [Kazusa Institute, www.kazusa.org.jp]; Nos. 7120, Nostoc PCC 7120 (Anabaena) [Kazusa]; Syn 8102, Synechococcus WH 8102 [DOE Joint Genome Institute, www.jgi.doe.gov]; G. viol, Gloeobacter violaceus [Kazusa].

The separate, but adjacent, single copy genes for the N- and C-terminal domains of the OCP in the genome of *Thermosynechococcus elongatus* (Table 2) suggest that the 35 kDa OCP is the product of a gene fusion event.

There are scattered reports of other cyanobacterial carotenoid proteins in the literature. A carotenoid protein was isolated from the cytoplasmic membrane fraction of *Synechocystis* PCC 6714 [7]. The apparent molecular mass of this protein, determined by SDS–PAGE, was estimated to be either 35 or 45 kDa depending on the solubilization temperature. This protein cross-reacted with antibodies prepared against a 46 kDa carotenoid protein isolated from *Anacystis nidulans* R2 [8]. The thylakoid membrane of *A. nidulans* R2 also contains a carotenoid protein of 42 kDa [9] but it shows no sequence homology to the OCP [10]. Fractionation of *A. nidulans* yields a red water-soluble carotenoid protein; this carotenoprotein contains zeaxanthin and is

a homodimer of two 23 kDa polypeptides [11]. A watersoluble zeaxanthin-binding protein complex of 58 and 56 kDa proteins was isolated from the cell envelope of Prochlorothrix hollandica [12]. As a group, carotenoidbinding proteins vary with regard to cellular location, detergent versus water solubility, and the type of associated carotenoid. Furthermore, care must be taken in interpretation based on molecular mass estimations for the OCPs; molecular weights determined by SDS-PAGE and size exclusion chromatography data have been known to vary significantly from that determined by mass spectrometry (D.W. Krogmann pers. comm.). One common feature noted for several of these carotenoid-binding proteins is that they increase in abundance in response to high light [9,10,12]. In the present discussion, the use of "OCP" is restricted to carotenoid proteins that are verifiably orthologs of the 35 kDa protein that corresponds to the slr1963 gene product of Syn 6803 [5].

2

ARTICLE IN PRESS





The three-dimensional structure of the OCP

The structure of the OCP of A. maxima has been determined to 2.1 Å resolution [13,14]. The asymmetric unit of the crystal contains an OCP dimer (Fig. 1) and the two molecules are essentially identical with few exceptions (noted below). Dimerization in the crystal buries 1411 \AA^2 of surface area, suggesting that the asymmetric unit corresponds to the dimer observed in solution (D. Krogmann, pers. comm.). Homodimerization is a distinctive feature of the OCP relative to known structures of lightharvesting proteins. The relative disposition of the carotenoids in the OCP dimer is also unique; the 3'-hydroxyechinenone molecules are nearly parallel. In contrast, the carotenoids in peridinin-chlorophyll protein [15]-the only known light-harvesting protein structure with multiple closely interacting carotenoids-show no obvious preference for parallel alignment. The carotenoids of the ring-shaped bacterial light-harvesting complexes (composed of nine heterodimeric units) are nearly parallel, but very widely spaced [16-20].



Fig. 1. The *A. maxima* OCP dimer. The N-terminal, all-helical domain is shown in dark gray, the C-terminal domain in light gray. The carotenoid is shown in white, in space-filling representation. Chloride ions are shown in black, and the sucrose molecule observed in one of the molecules in the asymmetric unit is shown in sticks. This figure and Figs. 3, 5, and 6 are prepared with Pymol [52].

C.A. Kerfeld | Archives of Biochemistry and Biophysics xxx (2004) xxx-xxx

The structure of the OCP consists of two domains (Fig. 1). The all- α -helical N-terminal domain resembles two four-helix bundles, however the helices forming each bundle are made up of discontinuous segments of the primary structure of the protein (Fig. 1, and Table 1) suggesting that the eight helices form an intact domain. The C-terminal domain resembles a NTF-2 (nuclear transport factor 2) domain as predicted by P-FAM analysis [21] of the OCP primary structure. NTF-2 domains are observed in a wide range of functionally distinct proteins including enzymes and transport proteins [22–26]. The NTF-2 fold is a mixed α/β fold, characterized by the formation of a hydrophobic pocket. In the OCP, the keto group of the 3'-hydroxyechinenone molecule is nestled into the core of the NTF-2 domain. The N- and C-terminal domains are joined by a long extended loop (residues 161-185). In one of the molecules of the dimer, a sucrose molecule (a component of the crystallization mother liquor) is nestled into a cavity formed by a cluster of highly conserved residues [14] in this linker region (Figs. 1 and 2A). The fructose moiety is in van der Waals distance of sidechain of Trp 279 adjacent to the 3'-hydroxyechinenone molecule (Fig. 3).

It is well known that the protein environment plays an important role in tuning the spectral characteristics of the pigment in pigment-protein complexes. This is dramatically demonstrated in the OCP (Fig. 4). The isolated carotenoid, 3'-hydroxyechinenone, appears yellow in organic solvents ($\lambda_{max} = 450$ nm) whereas it appears orange ($\lambda_{max} = 465$ and 495 nm) in the OCP. In the OCP, the 3'-hydroxyechinenone molecule has an all*trans* configuration with an average deviation from 180° of 16° and a radius of curvature of approximately 28 Å. The 3'-hydroxyechinenone molecule is approximately 26 Å in length and spans both domains of the OCP (Fig. 1).



Fig. 2. Surface and charge representation of the OCP. (A) The carotenoid molecule is yellow and shown in space filling representation. The sucrose molecule is yellow and rendered in sticks. The view in (B) is oriented 180° relative to the view shown in (A). Figure prepared with GRASP [53].



Fig. 3. Amino acid sidechains within 3.7 Å of the 3'-hydroxyechinenone molecule in the OCP. The sucrose molecule is shown in sticks. The chloride ion and nine water molecules (space-filling crosses) found in the carotenoid-binding cleft are also shown.



Fig. 4. Absorption spectra of the OCP (dotted line), the RCP (dashed line), and 3'-hydroxyechinenone in hexane (solid line) isolated from *A*. *maxima*.

The carotenoid binding cleft of the *A. maxima* OCP is lined with residues conserved in the primary structures of OCPs (Table 1) including Trp 290 and Tyr 203 which

ARTICLE IN PRESS

hydrogen bond to the keto oxygen atom of the 3'-hydroxyechinenone molecule. All residues within 3.7 Å of the carotenoid molecule are shown in Fig. 3. There are also eight water molecules surrounding the center of the carotenoid molecule at the interface between the two domains. The carotenoid binding cleft also contains a chloride ion hydrogen bonded to Thr 277 (Figs. 1 and 3), a conserved residue in the primary structure of the OCP.

A notable feature of the OCP structure is the presence of six Met residues (five absolutely conserved among the known primary structures of OCPs, Table 1), positioned so that their thioether groups extend toward the carotenoid molecule [14]. The importance of a Met residue in chromoprotein integrity has been demonstrated in bacteriophytochrome [27]. The unusual distribution of methionine residues in the OCP may indicate that these residues play an important yet undetermined role in its function. The flexibility of the Met sidechain as well as the polarizability of the sulfur atom could be important in binding the highly conjugated carotenoid molecule in the OCP. Furthermore, the oxidation of Met residues has been characterized in numerous biological systems and could be expected to occur in the context of the light reactions of photosynthesis. Conversion of the thioether moiety into Met sulfoxide or Met sulfone could influence the interaction of these sidechains with pigment. It is also possible that oxidation could affect the stability or alter the structure of the protein [28]. Methionine sidechains also have singlet oxygen quenching activity [29] which may contribute to the OCP's putative photoprotective function.

In the OCP, the carotenoid is almost entirely buried; only 3.4% of the surface of the 3'-hydroxyechinenone molecule is exposed to solvent. Solvent accessibility is restricted to two regions on the surface of the pigment molecule. The first is near the hydroxyl terminus of the carotenoid, which is wedged between the helical bundles (Fig. 1). The carotenoid is also slightly solvent accessible through the large elongated depression (Fig. 2B) formed by residues from both domains. The total volume of this cavity is 895 Å^3 , large enough for a substantial interaction with another protein. The cavity is continuous with the portion of the carotenoid binding pocket formed by the NTF-2 domain.

The red carotenoid protein

In the course of purification of OCPs from several cyanobacterial species, a red carotenoid protein (RCP) with distinctive optical properties (Fig. 4) was isolated [5,6]. N-terminal sequencing and mass spectrometry analysis indicates that this is a 16 kDa proteolytic fragment of the OCP [5]. This fragment is the result of removal of the first 15 and approximately the last 160

amino acid residues of the OCP. The proteolysis removes the entire C-terminal domain. Without concomitant structural changes (that are likely to occur in solution), this would result in the exposure of nearly half of the carotenoid to solvent (Fig. 5). The proteolysis of the OCP into an RCP also occurs under certain storage conditions. The 16kDa RCP elutes as a monomer in size-exclusion chromatography and has an isoelectric point (5.0) similar to that of the OCP (4.7) [6]. Aggregates of the 16kDa RCP induced by concentration or successive cycles of freeze-thawing appear orange and are retained by a YM 30 membrane [30], suggesting that native OCP spectral properties may be achieved through oligomerization of the RCP.

The observation of shorter OCP homologs in some cyanobacterial genomes (Table 2) resembling the RCP suggests that there may be smaller OCP-like proteins with altered spectral and oxygen quenching properties in these organisms. For example, in *Gloeobacter violaceus* and in both of the *Nostoc* genomes for which sequence data are available, in addition to full-length OCP genes, there are four paralogs that correspond to RCP-like proteins. These genomes also contain an open reading frame encoding a protein similar to the C-terminal, NTF2-like domain of the OCP. It is possible that different N-terminal domains may combine with the C-



Fig. 5. Hypothetical model of the 16 kDa RCP based on the known proteolysis sites and the structure of the OCP. The carotenoid is shown in the conformation it holds in the OCP form. The protein and pigment conformational changes that are likely to occur with proteolysis to the RCP are not yet known.

terminal domain and carotenoid to create spectrally distinct OCP-like proteins in these organisms. A paradigm for a modular assembly of carotenoid-containing subunits into distinct holoproteins occurs in crustacyanin, isolated from lobster carapace. Various combinations of five β -crustacyanin subunits (each ~20 kDa) combine to form dimers in which two carotenoids span the interface [31].

Spectral changes mimicking those in the 16 kDa RCP can be induced by acidification of the OCP. Mass spectrometry confirms that the acid-induced RCP has not been proteolyzed. The acid-induced RCP migrates similar to OCP under native PAGE conditions (Kerfeld, unpub.), suggesting no change in oligomerization in the conversion of the OCP to the acid-induced form of RCP. However, circular dichroism measurements indicate that it has altered secondary structure relative to the OCP. The acid-induced RCP is stable indefinitely at neutral pH; however, prolonged exposure in crystallization conditions similar to those used to crystallize the OCP (pH 8.5) results in orange crystals (Kerfeld, unpub.). There are precedents for reversible, low pH-induced spectral changes in photosensitive proteins. For example, green fluorescent protein undergoes a spectral shift in response to decreasing pH [32]. A pH-induced carotenoid spectral change in the context of photoprotective processes also occurs in the xanthophyll cycle of higher plants (reviewed in [33,34]), which is involved in the thermal dissipation of excess light energy in both Photosystem I and Photosystem II.

Putative functions of the OCP

The OCP is unusual in that its structural characterization precedes a precise understanding of its function. A photoprotective function for the OCP is supported by microarray data indicating that OCP transcript levels increase more than 600% upon transfer to high light [35]. Since their initial characterization, several specific photoprotective roles have been proposed for the OCP and other water-soluble carotenoid proteins; they include singlet oxygen quenching, carotenoid transport or light attenuation. Below, the putative functions of the OCP are considered in the context of the structural data.

Quenching singlet oxygen and triplet chlorophyll

Data from our laboratory show that the OCP is an effective quencher of singlet oxygen, with an activity comparable to that of copper superoxide dismutase [14]. This is a critical function in oxygen-evolving photosynthetic organisms which generate oxidizing molecules under a broad range of light intensities. Singlet oxygen quenching studies of the 16 kDa RCP show that it quenches singlet oxygen at a significantly faster rate

than the OCP [14], consistent with greater carotenoid accessibility. The methionine sidechains noted above may also play a significant role in the OCP and the RCP's efficacy as quenchers of singlet oxygen. However, the limited exposure of the carotenoid and the lack of evidence that the OCP localizes to the lumenal surface of the thylakoid membrane where these reactive oxygen species would be most concentrated make a physiological role in quenching singlet oxygen less likely.

Carotenoid-mediated photoprotective mechanisms also include quenching triplet chlorophyll and dissipating excited singlet chlorophyll in excess of the capacity of the electron transport chain (non-photochemical quenching). To carry out these quenching functions, the OCP would necessarily need to be able to dock to the photosynthetic apparatus. The numerous protuberances and cavities on the OCP surface could facilitate these interactions (Fig. 2).

Carotenoid transport

The OCP has also been suggested to be a carotenoid transport protein, perhaps involved in (re-)assembly of the photosynthetic apparatus. For example, Photosystem II components are also known to undergo rapid turnover in response to high light intensity. Levels of the transcripts for the Photosystem II D1 polypeptide (known to be vulnerable to UV-induced damage) and the OCP show a similar pattern of regulation in response to high light treatment [35]. Recently, phenotypic characterization of an OCP deletion mutant of *Syn* 6803 indicates that this strain is more sensitive to high light treatment (Diana Kirilovsky, pers. comm.) and to UV-B radiation (Imre Vass, pers, comm.).

Functioning as a transport protein, the OCP could shuttle carotenoids from their site of synthesis in the thylakoid membrane to other cellular destinations. In cyanobacteria carotenoids such as zeaxanthin, echinenone and carotenoid glycosides such as myxoxanthophyll are found in the outer membrane [36–38] and the cell wall [39-44]. High light or other environmental stresses results in an increase in the types and quantities of cyanobacterial carotenoids in these locations. Likewise, carotenoids are found in the cytoplasmic as well as the thylakoid membrane [45,46] where they have been implicated in regulating membrane fluidity. The possibility that OCP is a general carotenoid transport protein is supported by observations in our laboratory and that of our collaborator Roberto Bassi that the A. maxima apo-OCP can be reconstituted with a variety of nonnative carotenoids (Bassi, unpub.; Kerfeld, unpub.).

A role in transport may explain the conflicting results of early efforts to determine the cellular location of various carotenoid proteins, as well as the diversity of carotenoids associated with them [6–12]. Although speculative, several structural features of the OCP are consistent with a transport function. In a carotenoid transport function, the methionine residues near the carotenoid could be important for pigment binding and release. The sucrose molecule binds at a highly conserved region that structurally resembles an allosteric site. The binding site is between the two domains, an interface that presumably would need to open to release the carotenoid. In addition, circular dichroism measurements show that apo-OCP has significant secondary structure in solution, perhaps indicating that the protein adopts a stable folded form prior to binding or after release of pigment (Kerfeld, unpub.).

Component of photoprotective structures

The cell wall and the outer membrane of cyanobacteria are rich in protein as well as carotenoids [36–44]. Some of these proteins may be the OCP or its paralogs. For example, a protein of 37 kDa was observed in the carotenoid-rich outer membrane and cell wall fractions from *Synechococcus* PCC6307 (see Fig. 2 in [42]) but not further characterized. There are also proteins similar in size to the OCP homologs in the carotenoid-rich fraction of the outer membrane of *Synechocystis* PCC6714 [38]. Dissociation of the *P. hollandica* zeaxanthin-binding protein complex accompanied by staining for sugars indicates that it was composed of several proteins with molecular masses between 22 and 31 kDa [12].

In vivo studies of outer membrane carotenoids in *Synechocystis* PCC6714 indicate that the carotenoids are found at the periphery of the cell in a uniformly oriented array [47], suggestive of a photoprotective function. This spatial organization is suggested to be the result of the association of carotenoids with outer membrane proteins [47,48]. Packing of the OCP molecules in the crystals [13,14] is intriguing in that it suggests one way in which an array of aligned pigment molecules might be produced (Fig. 6). OCPs from some species are prone to

become covalently attached to some cellulose-based chromatography matrices [5], perhaps through chemistry similar to that required for association with the cell wall.

Variants of OCP and photoprotection

The photosynthetic apparatus of cyanobacteria is known to be structurally dynamic; the organisms alter the composition of their phycobilisomes in response to differences in the spectral composition of incident light [49]. Although speculative, the potential for modular assembly of OCPs in some cyanobacteria species could provide a means for tuning the response of the photosynthetic/photoprotective proteins to varying environmental conditions.

The OCP as a model system for study of carotenoidprotein interactions

The importance of carotenoids in the xanthophyll cycle and in anti-oxidant processes has stimulated efforts to manipulate the carotenoid content of different organisms. It is becoming apparent that biological systems are more versatile at synthesizing carotenoids than previously expected. For example, the introduction of an algal carotenoid biosynthetic gene into tobacco resulted in the accumulation of 3'-hydroxyechinenone [50], the cyanobacterial carotenoid found in the OCP. Furthermore, the importance of carotenoids as biological anti-oxidants has led to intensive efforts to express them in heterologous systems such as Escherichia coli [51]. To fully utilize this technology, it must be complemented by an understanding of how to integrate carotenoids into protein carriers which, for many applications, must be water-soluble. Spectroscopic and functional characterization as well as structural studies of the RCPs, apo-



Fig. 6. Packing of the *A. maxima* OCP molecules within the crystals used for structure determination. The C2 cell (traced) dimensions were a = 217 Å, b = 41 Å, c = 75 Å $\beta = 95.8^{\circ}$.

OCP, and site-directed mutants of the OCP will provide important information for manipulating carotenoid– protein content in other organisms.

Acknowledgments

This work was supported by the US Department of Agriculture (USDA 1999-01759). The author thanks David Krogmann, Todd Yeates, and Michael Sawaya for helpful discussions.

References

- H.A. Frank, R.J. Cogdell, Photochem. Photobiol. 63 (1996) 257– 264.
- [2] H.Y. Yamamoto, R. Bassi, in: D.R. Ort, C.F. Yocum (Eds.), Oxygenic Photosynthesis: The Light Reactions, Kluwer Academic, Dordrecht, The Netherlands, 1996, pp. 539–563.
- [3] D.R. Ort, Plant Physiol. 125 (2001) 29-32.
- [4] T.K. Holt, D.W. Krogmann, Biochim. Biophys. Acta 637 (1981) 408–414.
- [5] Y.P. Wu, D.W. Krogmann, Biochim. Biophys. Acta 1322 (1997) 1–7.
- [6] Y.P. Wu, Doctoral dissertation, Purdue University, 1996.
- [7] G.S. Bullerjahn, L.A. Sherman, J. Bacteriol. 167 (1986) 396-399.
- [8] G.S. Bullerjahn, L.A. Sherman, in: J. Biggens (Ed.), Progress in Photosynthesis Research, Martinus Nijhoff Publishers, Dordrecht, Netherlands, 1987, pp. II.1.145–148.
- [9] K. Masamoto, H.C. Riethman, L.A. Sherman, Plant Physiol. 84 (1987) 633–639.
- [10] K.J. Reddy, K. Masamoto, D.M. Sherman, L.A. Sherman, J. Bacteriol. 171 (1989) 3486–3493.
- [11] M. Diverse-Pierluissi, D.W. Krogmann, Biochim. Biophys. Acta 933 (1988) 372–377.
- [12] J.M. Engle, W. Burkhart, D.M. Sherman, G.S. Bullerjahn, Arch. Microbiol. 155 (1991) 453–458.
- [13] C.A. Kerfeld, Y.P. Wu, C. Chan, D.W. Krogmann, T.O. Yeates, Acta Cryst. D 53 (1997) 720–723.
- [14] C.A. Kerfeld, M.R. Sawaya, V. Brahmandam, D. Cascio, K.K. Ho, C.C. Trevithick-Sutton, D.W. Krogmann, T.O. Yeates, Structure 11 (2003) 55–65.
- [15] E. Hofmann, P.M. Wrench, F.P. Sharples, R.G. Hiller, W. Welte, K. Diedrichs, Science 272 (1996) 1788–11791.
- [16] G. McDermott, S.M. Prince, A.A. Freer, A.M. Hawthornthwaite-Lawless, M.Z. Papiz, R.J. Cogdell, N.W. Isaacs, Nature 374 (1995) 517–521.
- [17] A. Freer, S. Prince, K. Sauer, M. Papiz, A. Hawthornthwaite-Lawless, G. McDermott, R. Cogdell, N.W. Isaacs, Structure 4 (1996) 449–462.
- [18] J. Koepke, X. Hu, C. Muenke, K. Schulten, H. Michel, Structure 4 (1996) 581–597.
- [19] S.M. Prince, M. Papiz, A.A. Freer, G. McDermott, A. Hawthornthwaite-Lawless, R.J. Cogdell, N.W. Isaacs, J. Mol. Biol. 268 (1997) 412–423.
- [20] K. McCluskey, S.M. Prince, R.J. Cogdell, N.W. Isaacs, Biochemistry 40 (2001) 8793–8799.

- [21] A. Bateman, E. Birney, L. Cerruti, R. Durbin, L. Etwiller, S.R. Eddy, S. Griffiths-Jones, K.L. Howe, M. Marshall, E.L. Sonnhammer, Nucleic Acids Res. 30 (2002) 276–280.
- [22] T.L. Bullock, W.D. Clarkson, H.M. Kent, M. Stewart, J. Mol. Biol. 260 (1996) 422–431.
- [23] M. Stewart, H.M. Kent, A.J. McCoy, J. Mol. Biol. 277 (1998) 635–646.
- [24] S. Fribourg, I.C. Braun, E. Izaurralde, I. Conti, Mol. Cell 8 (2001) 645–656.
- [25] H.-S. Cho, G. Choi, K.Y. Choi, B.-H. Oh, Biochemistry 37 (1998) 8325–8330.
- [26] B. Kauppi, K. Lee, E. Carredano, R.E. Parales, D.T. Gibson, H. Eklund, S. Ramaswamy, Structure 6 (1998) 571–586.
- [27] S.J. Davis, A.V. Vener, R.D. Vierstra, Science 286 (1999) 2517– 2520.
- [28] E.R. Stadtman, J. Moskovitz, R.L. Levine, Antiox. Redox Signal. 5 (2003) 577–582.
- [29] I.B.C. Matheson, R.D. Etheridge, N.R. Kratowich, J. Lee, Photochem. Photobiol. 21 (1975) 165–171.
- [30] R. Knutson, MS thesis, Purdue University, 1998.
- [31] M. Cianci, P.J. Rizkallah, A. Olczak, J. Raftery, N. Chayen, P.F. Zagelsky, J. Helliwell, Proc. Natl. Acad. Sci. USA 99 (2002) 9795– 9800.
- [32] R.B. Robey, O. Ruiz, A.V.P. Santos, J. Ma, F. Kear, F.L.-J. Wang, C.-J. Li, A.A. Bernardo, J.A.L. Arruda, Biochemistry 37 (1998) 9894–9901.
- [33] B. Demmig-Adams, W.W. Adams, Annu. Rev. Plant Physiol. Plant Mol. Biol. 43 (1992) 599–626.
- [34] R. Bassi, S. Caffarri, Photosyn. Res. 64 (2000) 243-256.
- [35] Y. Hihara, A. Kamei, M. Kanehisa, A. Kaplan, M. Ikeuchi, Plant Cell 13 (2001) 793–806.
- [36] T. Omata, N. Murata, Arch. Microbiol. 139 (1984) 113-116.
- [37] U.J. Jurgens, J.R. Golecki, J. Weckesser, Arch. Microbiol. 142 (1985) 168–174.
- [38] U.J. Jurgens, J. Weckesser, J. Bacteriol. 164 (1985) 384-389.
- [39] M.J. Salton, Microbial Cell Walls, Wiley, New York.
- [40] N. Murata, N. Sato, T. Omata, T. Kuwabra, Plant Cell Physiol. 22 (1981) 855–866.
- [41] C. Resch, J. Gibson, J. Bacteriol. 155 (1983) 345-350.
- [42] D. Woitzik, J. Weckesser, U.J. Jurgens, J. Gen. Microbiol. 134 (1988) 619–627.
- [43] U.J. Jurgens, C. Martin, J. Weckesser, FEMS Microbiol. Lett. 65 (1989) 47–52.
- [44] R. Benz, H. Bohme, Biochim. Biophys. Acta 812 (1985) 286-292.
- [45] W.I. Gruszecki, in: H.A. Frank, A.J. Young, G. Britton, R.J. Cogdell (Eds.), The Photochemistry of Carotenoids, Kluwer Academic, Dordrecht, The Netherlands, 1999, pp. 363–379.
- [46] Z. Varkonyi, K. Masamoto, M. Debreczeny, O. Zsiros, B. Ughy, Z. Gombos, I. Domonkos, T. Farkas, H. Wada, B. Szalontai, Proc. Natl. Acad. Sci. USA 99 (2002) 2410–2415.
- [47] U.J. Jurgens, W. Mantele, Biochim. Biophys. Acta 1067 (1991) 208–212.
- [48] U.J. Jurgens, R. Benz, Z. Naturforsch. 44c (1989) 147-151.
- [49] D.M. Kehoe, A.R. Grossman, Science 273 (1996) 1409-1412.
- [50] V. Mann, M. Harker, I. Pecker, J. Hirschberg, Nat. Biotech. 18 (2000) 888–892.
- [51] P.C. Lee, A.Z.R. Momen, B.N. Mijts, C. Schmidt-Dannert, Chem. Biol. 10 (2003) 453–462.
- [52] W.L. Delano, 2002. Available from <www.pymol.org>.
- [53] A. Nicholls, R. Bharadwaj, B. Honig, Biophys. J. 64 (1993) A166.