### Bioinformatics-Based Identification of Chemosensory Proteins in African Malaria Mosquito, *Anopheles gambiae*

Zhengxi Li<sup>1\*</sup>, Zuorui Shen<sup>1</sup>, Jingjiang Zhou<sup>2</sup>, and Lin Field<sup>2</sup>

<sup>1</sup>Department of Entomology, China Agricultural University, Beijing 100094, China; <sup>2</sup>Biological Chemistry Division, Rothamsted Research, Harpenden, Herts. AL5 2JQ, UK.

Chemosensory proteins (CSPs) are identifiable by four spatially conserved Cysteine residues in their primary structure or by two disulfide bridges in their tertiary structure according to the previously identified olfactory specific-D related proteins. A genomics- and bioinformatics-based approach is taken in the present study to identify the putative CSPs in the malaria-carrying mosquito, *Anopheles* gambiae. The results show that five out of the nine annotated candidates are the most possible *Anopheles* CSPs of *A. gambiae*. This study lays the foundation for further functional identification of *Anopheles* CSPs, though all of these candidates need additional experimental verification.

Key words: chemosensory protein, proteomics, bioinformatics, olfaction, African malaria mosquito, Anopheles gambiae

## Introduction

Anopheles gambiae (A. gambiae) is the principal vector of malaria that afflicts more than 500 million people and causes more than 1 million deaths each year. Analysis of the whole genome of A. gambiae revealed strong evidence for about 14,000 protein-coding transcripts, which need further annotation and experimental verification (1).

Olfaction plays a key role in host selection of agricultural pests and disease vectors. Recent advancement in understanding the molecular mechanism of olfaction is the result of multidisciplinary research efforts by using a variety of model organisms including insects. The reception of pheromones and general odorants is mediated by specific neurons located in specialized cuticular sensilla in insects (2). Chemosensory neurons extend their dendrites to a lymphatic cavity, where the soluble and low molecular weight proteins that are supposed to transfer the hydrophobic odorants across the fluid barrier to the receptive dendritic membrane are contained. Molecular cloning and biochemical surveys of insect antennae have identified two abundant but unrelated families of small soluble proteins with proposed odorant transport function, that is, odorant-binding proteins

### \* Corresponding author.

E-mail: zxli@cau.edu.cn

(OBPs) and olfactory specific-D (OS-D) related proteins (3, 4). OS-D related proteins (average 13 kDa) were first identified by subtractive hybridisation experiments using antennae of *Drosophila melanogaster* (5, 6). Many OS-D homologues were subsequently identified based on sequence similarity in different insect orders (Table 1).

OS-D related proteins differ from OBPs in several aspects: they share no sequence similarity with OBPs and contain only four of the six spatially conserved Cysteine residues that are characterised by OBPs (4). Although two of the protein families are both represented by multiple genes within a given species, OS-D related proteins are more conserved than OBPs across evolution or between different phyla, with 40%-50% of identical residues even between most distant species. They have been identified in a variety of tissues, while most OBPs appear to be restricted to olfactory tissues. They are common in Orthopteroid (phasmid and grasshopper) and holometabolous (Lepidoptera and Diptera) insects (Table 1) and thus may be present throughout the Neoptera, while OBPs are only known within the holometabolous and hemipteroid lineages (4). There is no strong evidence for the physiological role of OS-Ds so far, and they may be involved in chemical communication and perception. To contrast with OBPs that are found in olfactory sensilla, the OS-D related proteins were designated as Chemosensory Proteins (CSPs; ref. 18, 23).

Order	Species	Protein name <sup>a</sup>	Length(a.a)	Accession No. <sup>b</sup>	References
Hymenoptera	Apis mellifera	ASP3c	130	AF481963	7, 8
Lepidoptera	$Cactoblastis\ cactorum$	CLP-1	130	U95046	9
	Manduca sexta	SAP1	105	AF117574	10
		SAP3	126	AF117585	10
		SAP2	127	AF117592	10
		SAP5	231	AF117594	10
		SAP4	127	AF117599	10
	Bombyx mori	BmorCSP2	120	AF509238	11
		BmorCSP1	127	AF509239	11
	Mamestra brassicae	CSP-MbraA1	112	AF211177	12
		CSP-MbraA2	112	AF211178	12
		CSP-MbraA3	112	AF211179	12
		CSP-MbraA4	112	AF211180	12
		CSP-MbraA5	112	AF211181	12
		CSP-MbraB1	108	AF211182	12
		CSP-MbraB2	108	AF211183	12
		CSP-MbraA6	128	AF255918	13
		CSP-MbraB3	108	AF255919	13
		CSP-MbraB4	108	AF255920	13
	Heliothis virescens	HvirCSP2	126	AY101511	14
		HvirCSP1	114	AY101512	14
		HvirCSP3	106	AY101513	14
	$Mamestra\ brassicae$	SAP	111	AY026760	unpublished
	Helicoverpa armigera	CSP-Harm	127	AF368375	unpublished
	Helicoverpa zea	CSP-Hzea	128	AF448448	unpublished
Diptera	D. melanogaster	A10	155	U05244	15
		m RH70879p	124	BT001865	unpublished
		PEBmeIII	158	U08281	16
	Anopheles gambiae	SAP-1	127	AF437891	17
Orthoptera	$Schistocerca\ gregaria$	CSP-sg1	109	AF070961	18
		CSP-sg2	109	AF070962	18
		CSP-sg3	103	AF070963	18
		CSP-sg4	109	AF070964	18
		CSP-sg5	109	AF070965	18
	$Locusta\ migratoria$	OS-D1	103	AJ251075	19
		OS-D2	120	AJ251076	19
		OS-D3	125	AJ251077	19
		OS-D4	125	AJ251078	19
		OS-D5	125	AJ251079	19
Phasmatodea	$Eury can tha\ calcarata$	CSP-ec1	107	AF139196	20
		CSP-ec2	102	AF139197	20
		CSP-ec3	107	AF139198	20
Dictyoptera	Periplaneta americana	p10	130	AF030340	$21.\ 22$

Table 1 Previously Identified Insect Chemosensory Proteins\*

\*GenBank (04/2003);

<sup>a</sup>Registered names in GenBank;

 $^{\rm b}{\rm GenBank}$  accession numbers.

Compared with the twenty-nine putative A. gambiae OBPs characterized for similarity to OBPs of Drosophila and other insects (24), no bioinformaticsbased annotation has been carried out to identify the CSP candidates of A. gambiae. The NMR (Nuclear Magnetic Resonance) solution structure of chemosensory protein Csp2 (1K19\_A) from moth Mamestra brassicae has been established (23), which is the best elucidated insect CSP and would be used as a model for homology modelling. We created an algorithm for identifying the conserved domains present in Anopheles putative CSPs through Perl programming.

# Results

## Conserved domain of insect CSP candidates

Exhaustive queries with all previously identified insect CSP sequences retrieved from GenBank (April 2003) and ClustalX multialignment resulted in an absolutely conserved structure for insect CSPs, that is, Cx(6,8)Cx(18)Cx(2)Cx(3) (Figure 1). We used a program developed by the authors through Perl programming to search the local database that contains the Fasta files of Anopheles gDNA (genomic DNA) and cDNA (complementary DNA) sequences downloaded from Ensembl Mosquito Genome Server for the identified pattern in their primary structure. Totally eight sequences were hit, including agCP10968. agCP11079, agCP11481, agCP11484, agCP11532, agCP11545, agCP6514, and agCP12965. The hits were in turn corroborated by BLAST searching in GenBank, and the prediction was made for their biochemical properties and secondary structure.

## CSP candidates corroborated by BLAST

The hits obtained through pattern searching were corroborated by BLAST in GenBank (http:// ncbi.nlm.nih.gov/). Six sequences (agCP10968, agCP 11079, agCP11481, agCP11484, agCP11532, and agCP11545) were found closely related to the previously identified CSPs, whereas two sequences (agCP6514 and agCP12965) matched no other proteins, so the CSP candidacy of agCP6514 and agCP12965 could not be excluded. The most interesting discovery was that a novel *Anopheles* CSP candidate (agCP11435) had been identified by BLAST searching, though the Expect (E) values are not very high (Table 2).

# Biochemical properties, prediction of secondary structure and ORFs of *Anopheles* CSP candidates

The cDNA sequences of the Anopheles Genome Project stored in GenBank and in the Ensembl Mosquito Genome Browser (ftp://ftp.ensembl.org/ pub/current\_mosquito/) had been annotated jointly by the privately funded Celera and EBI (http:// www.ensembl.org/Anopheles\_gambiae/). All the entries are the results from preliminary prediction. Even the positions of start codon for most of the annotated transcripts were not determined in the database. We have predicted the complete coding sequences (CDSs) and open reading frames (ORFs) for the Anopheles CSP candidates identified through pattern searching and BLAST, based on the genomic DNA sequences of the candidates. The redefined ORFs were listed in Table 3, along with the corresponding Celera IDs. GenBank IDs (#EAA), chromosomal locations, and scaffold numbers (#AAAB). Positions of the signal peptides, isoelectric points (pI) and hydrophobicity were also presented. The information contained in the table shows that all of the CSP candidates except agCP12965 and agCP6154 are located on one scaffold (AAAB01008964) of the chromosome 3R. This stimulating discovery may indicate that these olfactory genes were duplicated at some point of Anopheles evolution. Meanwhile, all of the CSP candidates have small molecular weights (<15 kDa in most cases) and are hydrophilic (<35% hydrophobic amino acids in most cases). Most of the candidates have a signal peptide, though no signal peptides have been found in two of the sequences, that is, agCP10968 and agCP12965. In fact, we have used several different tools to predict the ORFs of agCP10968, unfortunately no signal peptides have been found, which indicates that a sequencing error may have occurred in the Anopheles Genome Project.

Table 4 summarized the secondary structure prediction of *Anopheles* CSP candidates. The predictions showed that most of the candidates could be classified as all-alpha structure, with a high probability to form a globular domain. However, agCP10968 and agCP6514 did not appear to be globular.

	* 20 * 4	n +	60 ± 📕 80
CSP-ec1		7NYUNV	BEREGNERALESARK
CSP-ec3		7KYDNV	WENTFERERIFASMEETIGN : 21
CSP-sgl		BEKYT7KYDNV3	NEDETLANDREDNNWVKCLJEDGE : 35
CSP-sg2		BEKYT7MFDNW	AUDITAUDITUNNAVKENTEDGE : 33
CSP-894			LIGHT ANDREINKAWO THEDDE : 33
CSF-sg3		BEKYTTKYDNV	HIDERLANDELINKMACCLIENDD : 33
p10	MK-CVAVFVIVAVVALAEA	ARFRRDDKYT7KYDNI	LDEILASDRILANMHKELTEE : Se
08-D4	NQKCTLALLLACLVAA	TAAYT <mark>7</mark> KYDNII	DLEDWIHNDRIJKKAHE <mark>CHI</mark> SDSD : S:
OS-D5	MQKCTLALLLACLVAA	AAA <u>WT</u> ?KYDNII	DI DO HUDIU KKAHECU SDSD : 5:
OS-D1		A A A MARTINE	CAUDITHADRACKKAHECHASSSD : 21
08-02	MOALTLVLFALVAS	AAAVT7EYDNII	DADE DANKEDALKKNHERANSDED : 41
A10	MGQPGFRRAIGHFSLVVALMCTTCFQVEGLPHPPATSPS	DWMERMVEQAYDDKFDNV	DEDELLNQERLUINNIK PLEGT : 7
HvirCSPZ	MKFIVAVALLCLVAESWA	ASTYTERWONIS	VDEILESQRLIKA/VDCLIDR : 53
SAP4	MKMLLLV-ISCCVALFWAL	SASTYTDKWDNI3	WDEILESDRIMKGWVDCLIEDK : 52
PEBneIII	PLVVLGLVLVAA	BDKYT7KYDNI	WOBINKSDENEGNNEKSMODN : 43
ASP3c	MEVSTICIAIMOATAIAAAR-D	DESWDSKEDNT	AND THE STORN WARK THE
BmorCSP1	MKVLIVLSCVLVAVLA	DD-KYVDKYDKI	VIORI ENKINESANDENGK : 41
OLTP	MMK7SLVLLCCLAAVLARP	SD-TYTDKYDNI3	NICRILBNKRLIEANVNCVIDK : 52
CSP-Harn	MKVLLVLCLFAAAALA	DD-KYTDKYDNI3	ALDEILENKRLILANVNCVMER : 41
CSP-Hzea	MKVLLVLCLFAAAALA	DDDKYTDKYNNIK	ALDEILENKRLILAANNOWHER : 51
HV1FC8F3			APPEDIDENT AND A CONTRACT
CSP-MbraA2		BD-KWWDWDNT	HOELLANKEHIVANNEW ER : 33
CSP-NbraA5		BD-KYTONIS	LOETLANKELIVANNSVHER : 33
CSP-MbraA3		BD-KYTDKYDNI	LDEILANKELLVAAVNOVER : 33
CSP-MbraA6	MKFVLLLCVMVAAVVA	BD-KYTDKYDNIX	ALDEILANKRLIVAYVNCVHER : 41
CSP-MbraA4		BD-KYTDNIN	ARE TRANSFORMATION TO BE 3
SAP2	MKYLLVLCCVVAAVVC	DD-KWWDRWDRW	ANDSHIJAJERINIKGAVDEVIZR : 43
CSP-MbraB1		BEAHAADBAADAA	CHOST HGRINDARY PATKOTHD0: 3
CSP-NbraB2		ERAHYDDSYDSYD	DATE OF A CONTRACTOR CONTRACTOR CONTRACTOR
CSP-MbraB4		BEAHYTDRYDNVI	DINELOGIREDIVENCETUDE : 3
HvirCSP1	MALARP	DGAAYTDKYDNV	LDEILSNRRLLVP/VK2IIDQ : 40
SAP3	MK?FVALCLLSVVAV?LA	RPDHYTDRYDNV	LDEILDNHRVIVPVIKCILDQ : 52
SAP5	MKPVTAAILISLACMVQC	GKDMYTSRYDSN	WEDVIGNERERANIKOMPDB : 52
CSP-ec2		7 <u>RYDNV</u>	MAPSIN QHEBSANSHYNCHISL : 21
BmorCSP2	MKLLLVFLGLFLAVLA	QDKAEPIDeSF	ASDWISTERNIKSATKOWNQ: 41
SAPI	MKRLAVDCLEVDGALSAVE		MERIPERDERKALDERDK : 50
RH70879p	-IOLYTVSRSERMLLLNKNRVISLVVN	-FIFLIILISS VOATER	NUKRINIOVANSROIMGUGK : 61
		y ni	1 e6 y C6
000	* 100 * 120	* 10	40 = 160
CSP-ecl	* 100 * 120 CLAADAE HISKAI PDITENESAISEEKKAIVETTIVE	* 10 LIKNKDEI ESPKKKVDD	40 * 160 THKEKIERYIKQ : 93
CSP-ec1 CSP-ec3 CSP-sc1	100 * 120 GLADDASCIEKA PD. TENE AKOSKIKA VETTIVE RPOPDOCIMAL PD. TENE AKOSKIKA VETTIVE ANOVO	* 10 IKNKPET ESPKKNUDT IKNKPEV ESPKKNUDT	40 160 THK EKI ERVING 99 THK OTF DNLKG 99 DR EDESPER 90
CSP-ec1 CSP-ec3 CSP-sg1 CSP-sg2	* 100 * 120 GLGADDARDIKKAIPDI JENEGARGSEKIKA VETTIVE RECEDED OLIVEALA PDI JENEGARGSEKIKA VETTIVE ANOTVOKSI KAIPDI JENEGARGNEKE TEKI JEH ANOTADISI JEKAIPDI JENEGARGNEKE TEKI JEH	* 1 LIKNKDET ESPKKNUDT LIKNKPEV ESPKKNUDT LINHKPDI AQUKANUDE LINHKPDI AQUKANUDE	40 160 THK EKI ERYIKQ 99 THK QTF DNLKQ 99 DGT SK EDKSKE 100 DGT SK EDKSKE 100
CSP-ecl CSP-ec3 CSP-sgl CSP-sg2 CSP-sg2 CSP-sg4	* 100 * 120 CLEADARELIKATODA DENEGAROSEKTRA VETT VE RECORD OF LIKATODA DENEGAROSEKTRA VETT VE ANOTVO SKLADA DO JENEGARONEKE TKVVKH ANOTVO SLIKA PO JENEGARONEKE TKVVKH ANOTAD SLIKA PO JENEGARONEKE TKVVKH	* 1( LIKNKPEITESPEKKYDE LIKNKPEYTESPEKKYDE LINKKPDITADIKAKYDE LINKKPDITADIKAKYDE LINKKPDITADIKAKYDE LINKKPDYTADIKAKYDE	40 160 THK EKI ERYIKQ : 99 THK OTF DNLLKQ : 99 DGT SK EDKSKE : 100 DGT SK EDKSKE : 100 DGT SK EDKSKE : 100
CSP-ecl CSP-ec3 CSP-sgl CSP-sg2 CSP-sg4 CSP-sg5	100         120           GLGADAE         KAPDOI DENEARCESEKERA VETTUY           RECPORT         KAPDOI DENEARCESEKERA VETTVY           ANGTVD         KAVPOI SNEARCENERA VETTVY           ANGTVD         KAPDOI SNEARCENERA VETTVY           ANGTVD         KAPDOI SNEARCENERA VETTVY           SNEARCESEKERA         VETTVY           SNEARCESEKERA         VETTVY           ANGTVD         KAPDOI SNEARCENERARCESEKERA           SNEARCESEKERARCESEKERARCESEKERARCESEKER         SNEARCESEKERARCESEKER           SNEARCESEKER         SNEARCESEKER           SNEARCESEKER         SNEARCESEKER	* 14 L KNK DET ESPKK L KNK DEV ESPKK U NHKPDI AQUKA L NHKPDI AQUKA U NHKPDI AQUKA L NHKPDV AQUKA U NHKPDV AQUKA	40 160 THE EXI ERVIEQ
CSP-ecl CSP-ec3 CSP-sgl CSP-sg2 CSP-sg4 CSP-sg5 CSP-sg3	*     100     *     120       *     GLGADDAE     *     KAP DD     PENE AX CSEK KAP VETT VF       *     RPOPDOC     KAP DD     PENE AX CSEK KAP VETT VF       *     AN TVD     KAP DD     PENE AX CSEK KAP VETT VF       *     AN TVD     KAP DD     FENE AX CNEK KAP VETT VF       *     AN TVD     KAP DD     SNE AX CNEK KEPK       *     AN TAD     KAP DD     SNE AX CNEK KEPK       *     SN TAD     SV IPD     SNE AX CNEK KEPK       *     SN TAD     SV IPD     SNE AX CNEK KEPK       *     SN TAD     SV IPD     SNE AX CNEK KEPK	* 14 LIKNKDEI ESKKNUT LIKNKDEV ESPKNUT LINKKDI ADIKANUT LINKKDI ADIKANUT LINKKDV ADIKANUT LINKKDV ADIKANUT	40 160 THE EKI ERVIRO
CSP-ec1 CSP-ec3 CSP-sg1 CSP-sg2 CSP-sg4 CSP-sg5 CSP-sg3 p10	100 * 120 GLARDAR KAR DO PENE AL SERVICA VETT VF RPOPDO LA DO PENE AL SERVICA VETT VF AN TVO KA PD ENE AL SERVICA VETT VF AN TVO KA PD ENE AL SNOK E TK KH SN TAD SVIPD ENE AL SNEKKE TK KH SN TAD SVIPD ENE AL SNEKKE TK KH SN TAD SVIPD ENE AL SNEKKE TK KH SN TAD SVIPD FILE SNEKKE TK KH SN TAD SVIPD FILE SNEKKE TK KH	* 14 KNKDEI ESKK	40         160           THK EKI ERVIKQ
CSP-ec1 CSP-ec3 CSP-sg1 CSP-sg2 CSP-sg4 CSP-ag5 CSP-ag5 CSP-ag3 p10 OS-D4 OS-D4	100 * 120 GLADDAR IKAN DO PENE ALSEKTA VETT VF RPODDO LA DO PENE ALSEKTA VETT VF ANTVO KA PD ENE ALSEKTA VETT VF ANTAD KA PD ENE ALSEKTA VETT VF ANTAD SVID ENE ALSEKT K SN 7AD SVID ENE ALSEKTE FK SN 7AD SVID ENE ALSEKTE FK SN 7AD SVID PENE ALSEKTE FK SN 7AD SVID FN 7A SK 7A S	* 14 KNKPET ESPKK	40         160           THK EKI ERVIKQ
CSP-ec1 CSP-sg1 CSP-sg2 CSP-sg2 CSP-sg3 CSP-ag5 CSP-ag5 CSP-ag3 p10 OS-D4 OS-D5 OS-D1	*         100         *         120           *         GLADDAE         KAR DD TENE ALSEKTAA VETT VP         *         120           *         RP CPOT         DA DD TENE ALSEKTAA VETT VP         *         *         *         *         120           *         RP CPOT         DA DD TENE ALSEKTAA VETT VP         * <td< td=""><td>* 14 KNKDET ESPKK KNKPEVESPKK NHKPDT AQUKA UNHKPDT AQUKA UNHKPDT AQUKA UNHKPDT AQUKA UNHKPDT AQUKA UNHKPDT AQUKA UNHKPDT ES OK UNKKFPMESTOR UKSKPDL EPUEK</td><td>40 160 THE TEXT ERVIEQ</td></td<>	* 14 KNKDET ESPKK KNKPEVESPKK NHKPDT AQUKA UNHKPDT AQUKA UNHKPDT AQUKA UNHKPDT AQUKA UNHKPDT AQUKA UNHKPDT AQUKA UNHKPDT ES OK UNKKFPMESTOR UKSKPDL EPUEK	40 160 THE TEXT ERVIEQ
CSP-ecl CSP-sgl CSP-sgl CSP-sg2 CSP-sg4 CSP-sg5 CSP-sg3 pl0 OS-D4 OS-D4 OS-D1 OS-D1 OS-D3	* 100 * 120 GLADAE IKAP DU ENE ALSEKKA VETT VF RECEDO CLIAA DU ENE ALSEKKA VETT VF AN TVD KAVPD ENE ALSEKKA VETT VF AN TVD KAVPD ENE ALSEKKA VETT VF AN TAD KA PD ENE ALSEKKA VETT VF SN TAD SVIPD ENE ALSEKKE TE K SN TAD SVIPD TIME ALSEKKE ALSE NE AS TPD AL PD TIME ALSEKKA AS INF AS TPD AL PD TIME ALSEKKA AS INF SN TAD SVIPD ENE ALSEKKA SK SK SN TD SVIPD SVIPD ENE ALSEKKE TE K SN TAD SVIPD TIME ALSEKKE TE SK SN TAD SVIPD SVIPD SVIPD SVIPTE SK SK SK SK SN TAD SVIPD SVIPTE SK SK SK SK SK SN TAD SVIPTE SK SK SK SK SK SK SN TAD SVIPTE SK SK SK SK SK SK SK SN TAD SVIPTE SK SK SK SK SK SK SK SK SK SK ST SK	* 14 UKNKPET ESPKK 77 UNNKPUT AUTKA 77 UNNKPDT AUTKA 77 UNNKPDT AUTKA 77 UNNKPDT AUTKA 77 UNKKPD AUTKA 77 UNKKPN ESUK 77 UKKKPN ESUK 77 UKKKPL FUSK 77 UKKKPL FUSK 77 UKKKPL FUSK 77	40         160           THE EKI ERYIKQ
CSP-ecl CSP-ecl CSP-sgl CSP-sgl CSP-sg4 CSP-sg5 CSP-sg5 CSP-sg3 p10 OS-D4 OS-D4 OS-D5 OS-D1 OS-D3 OS-D3 OS-D2	* 100 * 120 GLADAR KARPD PREAMENT AND SERVICE VETTYF RRSPDDCLADAR DO PENE AND SERVICE VETTYF AND VOC KAPD ONE AND NOT ENTRY AND VOC KAPD SNE AND NOT ETTYF NAN TAD SVIPD SNE AND NEW TYF SN TAD SVIPD SNE AND NEW TYF TYP TAD SVIPD SNE AND NEW TYF TYP TAD SVIPD SNE AND NEW TYF ASTPD AND VOC SNE AND NEW TYF ASTPD AND VOC SNE AND NEW TYF TYP TAD SVIPD TYPE AND NEW TA ASTRE TYP TAD SVIPD TYPE AND NEW TA ASTRE	* 14 LIKNKPEI ESKKNYT LIKNKPEYESKKNYT LINKKPOI AQUKANYT LINKKPOI AQUKANYT LINKKPOV AQUKANYT LINKKPOV AQUKANYT LINKKPOV AQUKANYT LINKKPOT AQUKANYT LIKKKPI ESUSKYT LIKKKPI ESUSKYT LIKKKPI ESUSKYT LIKKKPI ESUSKYT	40         160           THE ERTIERTING         9           THE OFF         9           OFF SK. EDNILKQ         9           DOT SK. EDNILKQ         100           DGT SK. EDNIKE         100           DGT SK. EDNIKE         100           DGT SK. EDRIKE         122           IGS RO. DOBLKR         122           SGS RO. DOBLKR         100           NGT RC. GEELKK         122           SGS RO. DOBLKK         122           SGS RO. DOBLKR         122           SGS RO. DOBLKR         123
CSP-ec1 CSP-sg1 CSP-sg2 CSP-sg4 CSP-sg5 CSP-sg3 p10 OS-D4 OS-D4 OS-D5 OS-D1 OS-D3 OS-D2 A10	* 100 * 120 GLADAR KAR PD PREASES AS SERVICE VETTY REPORT SERVICE VETTY REPORT SERVICE VETTY REPORT SERVICE VETTY AN TVD KAR PD PREASES AS ONE OF THE VETTY AN TVD KAR PD PREASES AS ONE OF THE VETTY AN TAD SERVICE	* 14 LIKNKPEI ESKKNYYY LIKNKPEY ESYKKYYY LINHKPDI AGIKAYYY LINHKPDI AGIKAYYY LINHKPDY AGIKAYYY LINHKPDI AGIKAYYY LINHKPDI AGIKAYYY LINKKPH ESIGKYY LINKKPL EPISKYY LIKSKPDI PPISKYY LIKSKPDI PPISKYY LIKSKPDI PPISKYY LIKSKPDI PPISKYY LIKSKPDI PPISKYY	40         160           THK EKI ERVIKQ
CSD-ec1 CSD-ec3 CSP-sg2 CSP-sg2 CSD-ag5 CSD-ag5 CSD-ag5 CSD-ag5 OS-D4 OS-D4 OS-D5 OS-D1 OS-D2 A10 HvirCSF2	* 100 * 120 GLADDAE KAR DO BENEAS SERVA VETT VF RECEDO CLAD DO ENESAS SERVA VETT VF AN TVD LKA PD SNEAS SERVA VETT VF AN TVD KA PD SNEAS SERVA VETT VF AN TVD KA PD SNEAS SERVA VETT VF SNEAD SV PD SNEAS NEEKE TE SK SNEAD SV PD SNEAS NEEKE SK SK TPD SNEAS SK AS TPD AN PD TINE AC NEEKEA AS SK SS TPD AN DO TINE AS NEEKEN AS SK SS TPD SNEAS SNEEKEN AS SK SS TPD SNEAS SK SS TPD	* 1 KNK KPC ESPK KNK KPC ESPK NHK PDI AQUKA INHK PDI AQUKA	40         160           THK EKI ERVIKQ
CSP-ec1 CSP-sg1 CSP-sg2 CSP-sg2 CSP-sg4 CSP-ag5 CSP-ag5 CSP-ag5 OS-D4 OS-D4 OS-D4 OS-D1 OS-D1 OS-D3 OS-D2 A10 HvirCSF2 SAP4	*         100         *         120           :         GLADDAE         KALPDI         DENE ALSEKRA VETT         VP           :         GLADDAE         KALPDI         DENE ALSEKRA         VETT         VP           :         R.P. PPO         CLADDAE         KALPDI         DENE ALSEKRA         VETT         VP           :         AN TVD         KALPDI         DENE ALSEKRA         VETT         VP           :         AN TVD         KALPDI         SNE ALSEKRA         VETT         VP           :         AN TVD         KALPDI         SNE ALSEKRA         VETT         VP           :         AN TAD         KALPDI         SNE ALSEKRA         VETT         VP           :         :         SNE ALSEKRA         SNE SALSEKRA         VETT         SNE           :         :         :         :         SNE SALSEKRA         SNE         SNE         SNE           :         :         :         :         :         :         SNE	* 14 KNKET ESPK. 77 KNKEV ESPK. UNHKPDI AQUKA. UNHKPDI AQUKA. UNHKPDV AQUKA. UNHKPDV AQUKA. UNHKPDV AQUKA. UNHKPDV AQUKA. UNHKPDI AQUKA. UNHKPDI ESUK. KEKPDI FUSK. UKEKPDI TUSK. UKEKPDI TUSK. UKEKPDI FUSK. UKEKPDI FUSK. UNHKPDI KEVA. UNHKPI ESUK.	40         160           THE DEXI ERVINO         99           THE OT DULLGO         99           OCT SK EDNILKO         99           DOT SK EDNILKO         90           DOT SK EDNIKO         100           DOT SK EDNERE         100           DOT SK EDNERE         100           DOT SK EDRERE         127           TGS RO DOELKR         122           SGS RO DOELKR         122           SGS RO GELKK         122           SGF LK UPELLSA         122           SGT LK UPELLSA         122           SGT RI GEMKSK         122           SG NI OE KDKIBAVK         122           SG KU DE KDKIBAVK         122           SG KU DE KDKIBER         122
CSP-ec1 CSP-sg1 CSP-sg2 CSP-sg2 CSP-sg4 CSP-ag5 CSP-ag5 CSP-ag3 p10 OS-D4 OS-D4 OS-D4 OS-D1 OS-D3 OS-D2 A10 HvirCSF2 SAP4 PEBmeIII	* 100 * 120 GLADDAE I KAR DD ' ENE ALSEKKA VETT 'V R POPPO'CLIDA D' ENE ALSEKKA VETT 'V AN TVD 'L' KA PD ' ENE ALSEKKA VETT 'V AN TVD 'L' KA PD ' ENE ALSEKKA VETT 'V AN TVD 'L' KA PD ' ENE ALSEKKA VETT 'V AN TAD 'L' KA PD ' ENE ALSEKKE 'K' KH SN TAD 'L' KA 'D' 'SNE ALSEKKE 'K' KH SN TAD 'L' KA 'D' 'SNE ALSEKKE 'K' KH SN TAD 'L' KA 'D' 'SNE ALSEKKE 'K' KH 'N'TD' 'L' KA 'D' 'SNE ALSEKKE 'K' KH 'N'TD' 'L' KA 'D' 'SNE ALSEKKE 'K' KH 'SN TAD 'L' KA 'D' 'SNE ALSEKKE 'SN' KH 'SN TAD 'L' KA 'D' 'SNE ALSEKKE 'SN' KH 'SN TAD 'L' KA 'D' 'SNE ALSEKKE 'SN' KH 'SN' 'SNE 'SNE 'SNE 'SNE 'SNE 'SNE 'SNE	* 14 KNK KPU ESPK UNHKPDI AQUKA UNHKPDI AQUKA UNHKPDI AQUKA UNHKPDI AQUKA UNHKPDI AQUKA UNHKPDI AQUKA UNHKPDI QUKA UNHKPDI QUKA UNHKPDI ESUK UKSKPI ESUK	40         160           THE EKI ERVIKO         99           THE OF SK. EDNIKO         99           OCT SK. EDNIKO         90           DOT SK. EDNIKO         100           DOT SK. EDNIKO         100           DOT SK. EDNIKE         100           DCT SK. EDNIKE         100           DCT SK. EDRIKE         100           DCT SK. EDRIKE         100           DCT SK. EDRIKE         100           DCT SK. EDRIKE         121           FGS RO. DOELKR         122           FGS RO. DOELKR         122           SCS RO. GPELKK         122           EGT LK HPELLSA         121           SCT RI GEMKIK         151           NNI OD KDKIEAVK         122           NNI OE KDKIEAVK         122           EEI IK. ROFGEDGOVERIEF         122           EEI IK. ROFGEDGOVERIEF         123
CSP-ecl CSP-sgl CSP-sgl CSP-sg2 CSP-sg4 CSP-sg3 pl0 OS-D4 OS-D5 OS-D1 OS-D2 A10 HvirCSF2 SAP4 PEBmeIII SAP-1 ASF3c	* 100 * 120 GLADDAE KAR PD PREAMSENTA VETTVF RRSPPOCLEDA PD PENEAMSENTA VETTVF ANDVD KERAPD POPERS AND KERAPETVF ANDVD KERAPD SNEAM NORKE TEN KH SN TAD STAN AND POPERS AND KERE TEN KH SN TAD STAN AND TIME AND KERE TEN KH SN TAD STAN AND POPERS AND KERE TEN KH SN TAD STAN AND THE AND THE AND KERE TEN KH SN TAD STAN AND THE AND THE AND KERE TEN KH SN TAD STAN AND THE AND THE AND KERE TEN KH SN TAD STAN AND THE STAN SN TAD STAN AND THE AND TH	* 14 UKNKPET ESPKN V 77 UNKKPDT AUTKA V 77 UKKKPM ESUC 77 UKKKPDT EPTSK V 77 UKKKPDT EPTSK V 77 UKKKPDT EPTSK V 77 UKKKPDT EPTSK V 77 UKKKPDT EFTSK V 77 UKKKPDT EFTSK V 77 UKKKPT EFTSK V 77 UKKKPT DOLK V 77 UKKKV 77 UKKKV 77 UKKKV 77 UKKKV 77 UKKKV 77 UKKKV 77 UKKV 77	40         160           THE EKI ERYIKO         99           THE OF SK. EDNILKO         99           OCT SK. EDNILKO         90           DGT SK. EDNILKO         100           DGT SK. EDNIKE         100           DGT SK. EDNIKE         100           DGT SK. EDNIKE         100           DGT SK. EDNIKE         100           DGT SK. EDRIKE         100           DGT SK. EDRIKE         100           DGT SK. EDRIKE         121           TGS RO DOELKR         122           SGS RO GPELKK         100           NGT RO GERLK         122           EGT KI DEHKK         122           EGT RI DEMKSK         122           EGT RI DEMKSK         122           EGT RI DEMKSK         122           EGT RI DEMKSK         122           EGT IK ROPRGEDQOVRRIRF         122           EEI IK ROPRGEDQOVRRIFF         122           EEI IK ROPRGEDQOVRRIFF         122           ENK LE RG-QAQKE         122           ENK KLE RG-GAQKE         122           ENK KLE RG-GACKE         122           ENK KLE RG-GACKE         122
CSP-ecl CSP-ecl CSP-sgl CSP-sgl CSP-sg2 CSP-sg3 pl0 OS-D4 OS-D5 OS-D1 OS-D3 OS-D3 OS-D2 A10 HvirCSF2 SAP4 PEBmeIII SAP-1 ASP3C HmorCSP1	* 100 * 120 GLADAR KAR PD PREAMENT AND SERVICE VETTYP RECORD CLEAR AND PD PENE AND SERVICE VETTYP AND YD CLEAR AND PD PENE AND SERVICE VETTYP AND YD CLEAR AND PD PENE AND NOT E THE VEH SN TAD SLEAR PD PENE AND NEW THE THE VEH SN TAD SLEAR PD POPENE AND NEW THE THE VEH SN TAD SLEAR PD POPENE AND NEW THE THE VEH SN TAD SLEAR PD POPENE AND NEW THE THE VEH SN TAD SLEAR PD POPENE AND NEW THE THE VEH SN TAD SLEAR PD POPENE AND NEW THE THE VEH SN TAD SLEAR PD POPENE AND NEW THE THE VEH SN TAD SLEAR PD POPENE AND NEW THE THE VEH SN TAD SLEAR PD POPENE AND NEW THE THE VEH SN TAD SLEAR PD POPENE AND NEW THE THE VEH SN TAD SLEAR PD POPENE AND NEW THE THE VEH SN TAD SLEAR PD POPENE AND NEW THE THE VEH SN TAD SLEAR PD POPENE AND NEW THE THE VEH SN TAD SLEAR PD POPENE AND NEW THE AND NEW THE THE VEH SN THE SLEAR PD POPENE AND NEW THE AND NEW THE AND NEW THE AND NEW THE AS TED SLEAR PD POPENE AND NEW THE AND NEW THE AND NEW THE SECTED AND SLEAR PD POPENE AND NEW THE AND NEW THE AND NEW THE GROUPD SLEAR ST PD POPENE AND NEW THE AND NEW THE AND NEW THE GROUPD SLEAR ST PD POPENE AND NEW THE AND NEW THE AND NEW THE GROUPD SLEAR ST PD POPENE AND NEW THE AND NEW THE AND NEW THE GROUPD SLEAR ST PD POPENE AND NEW THE AND NEW THE AND NEW THE GROUPD SLEAR ST PD POPENE AND NEW THE AND NEW THE AND NEW THE GROUPD SLEAR ST PD POPENE AND NEW THE AND NEW THE AND NEW THE SLEAR ST PD THE NEW THE SLEAR ST POPENE SCOTENT AND SCOTENT AND NEW THE SLEAR ST PD THE AND NEW THE SLEAR ST PD THE NEW THE SLEAR ST PD THE SLEAR ST PD THE NEW THE SLEAR ST PD THE SLEAR ST POPENE SCOTENT AND SCOTENT AND NEW THE SLEAR ST PD T	* 14 LIKNKPEI ESKKNJT LIKNKPEV ESPKNJT LINKKPOT AOTKANJE LINKKPOT AOTKANJE LINKKPOT AOTKANJE LINKKPOT AOTKANJE LINKKPOT AOTKANJE LIKKVPI ESTOR LIKKVPI ESTOR LIKKVPI ESTOR LIKKVPI ESTOR LIKKVPI ESTA LIKKVPI ESTA L	40         160           THE EKI ERVIRO         99           THE OF INLEQ         99           OGT SK EDNILKO         90           DGT SK EDNILKO         100           DGT SK EDNIKE         100           DGT SK EDNEKE         100           DGT SK EDREKE         122           FGS RO DOELKR         122           SGS RO DOELKR         122           SGS RO DELKR         122           SGF RI DELKK         122           EGT RI DEMKSK         122           EGT RI DEMKSK         122           EGT RI DEMKSK         122           EST IK ROPECOGUQURRIFF         122           ENK LE RG-QAQKE         122           ENK LE RGRAKG         123           OKK RW EBRAKKG         124
CSP-ec1 CSP-sg1 CSP-sg2 CSP-sg2 CSP-sg3 pl0 OS-D4 OS-D5 OS-D1 OS-D2 A10 HvirCSF2 SAP4 PEBmeIIII SAP-1 ASP3c EmorCSF1 OlfP	* 100 * 120 GLADDAE LKAR DDI ENE ALSEKTA VETT VF RECEDO CLOAD DDI ENE ALSEKTA VETT VF AN TVD CLAAD PDI ENE ALSEKTA VETT VF AN TVD CLAAD PDI ENE ALSEKTA VETT VF AN TVD CLAAD PDI ENE ALSEKTA VETT VF SN TAD CLAAD PDI ENE ALSEKTA VETT VF AS TPD CLAAD PDI THE ALSEKTA AS VETT AS TPD CLAAD PDI THE ALSEKTA AS VETT AS TPD CLAAD PDI THE ALSEKTA AS VETT AS TPD CLAAD PDI THE ALSEKTA AS VETT ST TT TAD CLAAD PDI THE ALSEKTA AS VETT SK TPD CLAAD PDI THE ALSEKTA AS SK TRM SK TPD CLAAD PDI THE ALSE SK TRM SK TRM SK TPD CLAAD PDI THE ALSEKTA AS SK TRM SK TPD CLAAD PDI THE ALSE SK TRM SK TPD CLAAD PDI THE ALSE SK TRM SK TPD CLAAD PDI THE ALSE SK TRM	* 14 KNK KEY ESPKK KNK KEY ESPKK NHK KDI AQUKA UNHK KDI AQUKA UNHK KDI AQUKA UNHK KDI AQUKA UNHK KDI AQUKA UNHK KPU ESOK UNHK KPU ES	40         160           THE DEXI ERVING         99           THE OTF DNLLKQ         99           DOGT SK. EDKEKE         100           DGT SK. EDKEKE         100           DGT SK. EDKEKE         100           DGT SK. EDREKE         121           TGS RO. DOBLKE         122           SGS RO. GPELKE         122           SGF LK         122           SGF LK         122           SGF LK         122           SGF RO. CEELK         122           SGF RO. DOKKENE         122           SGF RO. CEELSA         122           SGF RO. COCKRENE         122           SGF RO. COCKRENE         122           SGF RO. COCKRENE         122           SGF RO. COCKRENE         122
CSP-ec1 CSP-sg1 CSP-sg2 CSP-sg3 CSP-sg3 pl0 OS-D4 OS-D4 OS-D1 OS-D1 OS-D3 OS-D2 A10 HwirCSF2 SAP4 PEBmeIII SAP-1 ASP3C HmorCSF1 Olfp CSP-Harm	* 100 * 120 GLADDAE LKAR DD PENE ALSEKNA VETT VF RECEPCIANA DD FENE ALSEKNA VETT VF AN TVD LKAPD POLENE ALSEKNA VETT VF AN TVD LKAPD SNE ALSEN ALSEN VETT VF AN TVD LKAPD SNE ALSEN ALSEN VETT VF SNEAD SNE ALSEN	* 14 KNKDEI ESPKKYT NHKPDI AQUKAYT NHKPDI AQUKAYT NHKPDI AQUKAYT NHKPDI AQUKAYT NHKPDI AQUKAYT NHKPDI AQUKAYT NHKPDI AQUKAYT NHKPDI AQUKAYT NHKPDI ESUKYT KEKPDI EFUSKYT KEKPDI FUSKYT KEKPDI FUSKYT KEKPDI FUSKYT KEKPDI FUSKYT NKROL KUYCAYT NKROL KEYAAYT NKROL KEYAAYT NKROL KEYAAYT NKROL KEYAAYT NKROL KEYAAYT	40         160           THE EKI ERYIKQ
CSP-ec1 CSP-sg1 CSP-sg2 CSP-sg2 CSP-sg5 CSP-sg5 CSP-sg5 CSP-sg3 P10 OS-D4 OS-D5 OS-D1 OS-D3 OS-D2 A10 HvirCSF2 SAP4 PEBmeIIII SAP-1 ASF30 HmorCSF1 OlfP CSP-Harm CSP-Harm	* 100 * 120 GLADDAE LKAR DD PENE ALSEKKA VETT VF RESPONCE LAAD DD FINE ALSEKKA VETT VF AN TVD L KAP DD FINE ALSEKKA VETT VF AN TVD L KAP DD FINE ALSEKKA VETT VF AN TVD L KAP DD FINE ALSEKKA VETT VF AN TAD STAR DD FINE ALSEKKA VETT VF AN TAD STAR DD FINE ALSEKKA VETT VF SN TAD STAR DD FINE ALSEKKE TELSK SN TAD STAR DD FINE ALSEKKE TELSK AS TPD ALSEK DD FINE ALSEKKE TELSK AS TPD ALSE SN SD FINE ALSEKKE TELSK AS TPD ALSE SN STAR DD FINE ALSEKKE TELSK AS TPD ALSE SN STAR STAR ALSE SN ST TO ALSE STAR DD FINE ALSEKKE SN SK ST TO ALSE STAR DD FINE ALSEKKE SN SK ST TO ALSE STAR STAR STAR STAR ST TO ALSE STAR STAR STAR ST TO ALSE STAR STAR STAR STAR ST TO ALSE ST TO STAR STAR STAR STAR ST TO ALSE STAR STAR STAR STAR ST TO ALSE STAR STAR STAR STAR ST TO ALSE ST TO ST TAR STAR STAR STAR ST TO ALSE ST TO ST TAR STAR STAR STAR ST TO ALSE ST TO ST TAR STAR STAR STAR ST TO ALSE ST ST TO ST TAR STAR STAR STAR ST	* 14 UKNKDEI ESPKNYJT UNHKEDI AQUKAYJT UNHKEDI AQUKAYJT UNHKEDI AQUKAYJT UNHKEDV QUKAYJT UNHKEDV QUKAYJT UNKKEDI EPUEKYJT UKKEDI EPUEKYJT UKKEDI EPUEKYJT UKKEDI EPUEKYJT UKKEDI EPUEKYJT UKKEDI KEJANJT UKKEDI KEJANJT UKKEDI KEJANJT UNKKOJ KEJANJT UNKROJ DVUCKEJA	40         160           THE EKI ERYIKO         99           THE OF SK. EDKEKE         90           DOT SK. EDKEKE         100           DOT SK. EDKEKE         100           DGT SK. EDKEKE         100           DGT SK. EDKEKE         100           DGT SK. EDREKE         121           TGS RO. DOELKE         122           TGS RO. DOELKE         122           SCS RO. GPELKE         122           SCG K. UPELLSA         122           SCT RI GEMEKE         122           SCT RI GEMEKE         122           SCT RI GEMEKE         122           SCT RI GEMEKE         122           SCT RI GEMEKER
CSP-ecl CSP-sgl CSP-sgl CSP-sg2 CSP-sg3 pl0 OS-D4 OS-D5 OS-D1 OS-D2 A10 HvirCSP2 SAP4 PEBmeIII SAP-1 ASP3c EmorCSP1 OlfP CSP-Harm CSP-Harm CSP-Harm	* 100 * 120 GLADDAS KAR PD PLENE ALSEKTA VETT VF RRSPPD CLEAR DA PD PLENE ALSEKTA VETT VF AND VD KLEAR PD SNE ALSEKTA VETT VF AND VD KLEAR PD SNE ALSEKTA VETT VF AND VD KLEAR PD SNE ALSEKTA VETT VF SN 7AD SLEAR PD SNE ALSEKTE TELSKE SN 7AD SLEAR PD TENE ALSEKTE TELSKE AS 7PD SLEAR PD TINE ALSEKTE ALSE NE AS 7PD SLEAR PD TINE ALSEKTE ALSE NE AS 7PD SLEAR PD TINE ALSEKTE ALSE NE AS 7PD SLEAR PD TINE ALSEKTE ALSE NE GE 7PD ALSE TELSKE GE 7PD AL	* 14 L KNK KPE I ESKK VOT L KNK KPU ESPK NHK PDI AQUKA VOT L NHK PDI AQUKA VOT L NHK PDV QUKA VOT L NHK PDV QUKA VOT L NHK PDI AQUKA VOT L KEK PDL EPUSK VOT KEK PDL FUSK VOT L KEK PL POT AN VOT L KNEL FUSK VOT AN VOT VOT AN VOT AN VOT VOT AN VOT AN VOT AN VOT AN VOT VOT AN VOT AN VOT AN VOT AN VOT AN VOT VOT AN VOT A	40         160           THE EKI ERYIKO         99           THE OF SK. EDNILKO         99           OCT SK. EDNILKO         90           DGT SK. EDNILKO         100           DGT SK. EDNIKE         100           DGT SK. EDREKE         100           DGT SK. EDREKE         100           DGT SK. EDREKE         120           DGT SK. EDREKE         121           TGS RO DOELKR         122           SGS RO DOELKR         122           SGS RO GPELKK         121           SGT LK IPELLSA         122           SGT LK IPELLSA         122           SGT KK IPELLSA         122           SGT KK IPELLSA         122           SGT KK IPELSA         122           SGT KK IPELSA         122           SGT KK ROP RE RACKEL         122           SGT KK EDRAKAKG         122           SGT KK EDRAKAKG         122           SGK RK EDRAKAKG         122           SGK RK EDRAKAKG         122           SGK RK EDRA
CSD-ecl CSP-sgl CSP-sg2 CSP-sg2 CSP-sg3 CSD-ag3 pl0 OS-D4 OS-D5 OS-D1 OS-D2 OS-D2 Al0 HvirCSP2 SAP4 PEBmeIII SAP-1 ASP3c HmorCSP1 OlfP CSD-Harm CSP-MbraA1 CSP-MbraA1	* 100 * 120 GLADDAE LKAR DOT ENE ALSEKTA VETT VF RP CPO CLOAD DO ENE ALSEKTA VETT VF AN TVD LKA PD SNE ALSEKTA VETT VF AN TVD LKA PD SNE ALSEKTA VETT VF AN TVD LKA PD SNE ALSEKTA VETT VF SNEAD SV PD SNE ALSEKTE TELSK SNEAD SV PD SNE ALSEKTER ALSE SV SNE SNE SNE SNE SNE SNE SNE SNE SK SNE SNE SNE SNE SNE SNE SNE SNE SNE SNE	* 1 KNK KEY ESPKK NHKPDI AQUKA NHKPDI AQUKA NHKPI A	40         160           THE EKI ERVIRO         99           THE OF SK. EDNIKO         99           OGT SK. EDNIKO         90           DGT SK. EDNIKO         100           DGT SK. EDNIKE         100           DGT SK. EDNIKE         100           DGT SK. EDNIKE         100           DGT SK. EDRIKE         100           DGT SK. EDRIKE         100           DGT SK. EDRIKE         100           DGT SK. EDRIKE         121           TGS RO. DOBLKE         122           SGS RO. DOBLKE         122           SGT KI UPELLSA         121           NMI OD. KDKIEAVK         122           SGT KI ROPELKK         122           EGT KI QEMKSK         122           DOK KDKIEAVK         122           EET IK ROPRGSDOGVRRIRF         122           DOK RK EDRAAKAG         122           DOK RK EDRAAKAG         122           KGD RK & E
CSP-ec1 CSP-sg1 CSP-sg2 CSP-sg2 CSP-sg3 pl0 OS-D4 OS-D4 OS-D3 OS-D3 OS-D3 OS-D3 OS-D3 OS-D3 OS-D3 OS-D3 OS-D3 OS-D4 PEBneIII SAP-1 ASP3C EmorCSP1 OIFP CSP-H2ea HwirCSP3 CSP-MbraA1 CSP-MbraA5	* 100 * 120 GLADDAE LKAR DDI ENE ALSEKKA VETT VF RECEDO CLIDA DDI ENE ALSEKKA VETT VF AN TVD CLIKA PDI ENE ALSEKKA VETT VF AN TAD SLIKA PDI ENE ALSEKKE TELKIK SN TAD SLIKA PDI ENE ALSEKKE TELKIK SN TAD SLIKA PDI ENE ALSEKKE TELKIK TN TDD SLIKA PDI ENE ALSEKKE TELKIK GKOTPD SLIKA PDI TNE ALSEKKE TELKIK AS TPD SLIKA PDI TNE ALSEKKA AS TRF AS TPD SLIKA PDI TNE ALSEKKA AS TRF AS TPD SLIKA PDI TNE ALSEKKA AS TRF AS TPD SLIKA PDI TNE ALSEKKA AS TRF GETTER AN TRF ALSEKKE TELKIK GKOTPD SLIKA PDI TNE ALSEKKA AS TRF GETTER ALSEKKE TELKIK GROTD SLIKA PDI TNE ALSEKKA AS TRF GROTD SLIKA PDI TNE ALSEKKA PI AS TRF GROTD SLIKA PDI TNE ALSEKKA PI AS TRF GROTD SLIKA PDI TNE ALSEKKA PI AS TRF GROTD SLIKA PI PI DI TRF SLIKA STR PI AS TRF GROTAE NI RV PDI ATTOK CONNERTING SKAY TRF GROTAE NI RV PI ATTOK CONNERTING FRANK FROM GROTAE SLIKA PDI TRF SLIKA PI AS TRF GROTAE NI RV PI ATTOK CONNERTER FRANK FROM GROTAE STR PI DI BETGE CON AS KAY THE GROTAE STR PI DI BETGE CON AS KAY THE GROTE SE SLIKA PDI ENG SCARA SKAY THE GROTE SE SLIKA PDI ENG SCARA SKAY THE GROTAE STR FRANK FROM FRANK FROM FRANK FROM GROTAE STR FRANK FROM FROM GROTAE STR FRANK FROM GROTAE STR FROM GROTAE STR FROM GROTAE STR FROM GROTAE STR FROM GROTAE STR FROM GROTAE STR FROM GROTAE	* 14 CKNKEEI ESPKK 77 CKNKEPC ESPKK 77 CNHKEDI AQUKA 75 CNHKEDI AQUKA 75 CNHKEDV AQUKA 75 CNHKEDV AQUKA 75 CNHKEDV AQUKA 75 CNHKEDI ESUK 75 CNHKEDI ES	40         160           THE EKI ERYIKQ
CSP-ec1 CSP-sg1 CSP-sg2 CSP-sg2 CSP-sg3 pl0 OS-D4 OS-D4 OS-D5 OS-D1 OS-D3 OS-D2 A10 HvirCSF2 SAP4 PEBneIII SAP-1 ASP3c EmorCSF1 OlfP CSP-Hzea HvirCSF3 CSP-MbraA5 CSP-MbraA3	* 100 * 120 GLADDAE LKAR DD PENE ALSEKTA VETT VF RECEPCIONAL KAUPD PENE ALSEKTA VETT VF AND VD SLIKA PD PENE ALSEKTA VETT VF AND VD SLIKA PD PENE ALSEKTA VETT VF AND VD SLIKA PD PENE ALSEKTA VETT VF AND AD SLIKA PD PENE ALSEKTA VETT VF SN CAD SLIKA PD PENE ALSEKTE TELLEN SN CAD SLIKA PD FINE ALSEKTE TELLEN SN CAD SLIKA PD TINE ACONENTE ALSEKTE AS CPD SLIKA PD TINE ACONENTA AS VIE AS CPD SLIKA PD TINE ACONENTA AS VIE AS CPD SLIKA PD TINE ACONENTA AS VIE AS CPD SLIKA PD TINE ACONENTA AS VIE SN CAD SLIKA PD TINE ACO	* 14 CKNKET ESKNYTT CNKKEVESSKYTT CNKKEVESSKYTT CNKKEVESSKYTT CNKKEVT AUTRA CNKKEVT AUTRA CKKELSSKYTT CKKELSSKYT	40         160           THE EKI ERYIKQ
CSP-ecl CSP-sgl CSP-sgl CSP-sg2 CSP-sg3 CSP-sg3 CSP-sg3 OS-D4 OS-D5 OS-D1 OS-D3 OS-D2 A10 HvirCSF2 SAP4 PEBmeIII SAP-1 ASF3c HmorCSF1 OlfP CSP-Harm CSP-Harm CSP-Harm CSP-HarA3 CSP-MbraA3 CSP-MbraA3	* 100 * 120 GLADDAE LKAR DD PINE ALSEKNAA VETT VF RRSPDOCLINA DD PINE ALSEKNAA VETT VF ANDVOCLINKA DD PINE ALSEKNAA VETT VF ANDVOCLIKA PD PINE ALSEKNA VETT VF ANDAD SIKA DD ENE ALSEKNA VETT VF ANDAD SIKA DD ENE ALSEKNA VETT VF SN GAD SIKA DD ENE ALSEKNE TELVEN SN GAD SIKA DD ENE ALSEKNE TELVEN SN GAD SIKA DD ENE ALSEKNE TELVEN TN GDD SIKA DD ENE ALSEKNE TELVEN GK TPD SIKA DD ENE ALSEKNE TELVEN AS TPD SIKA DD TINE ALSEKNE ALSEKNE AS TPD SIKA DD TINE ALSEKNEN AS TPD SIKA DD TINE ALSENE ALSEN AS TPD SIKA DD TINE ALSENE ALSEN GRUTPE SIKA DD TINE ALSEN ALSEN ALSEN G	* 14 UKNK DEI ESKK VOT UKNK DEV ESSK VOT UNKK DI AUKA VOT UNKK DI AUKA VOT UNKK DI AUKA VOT UNKK DI AUKA VOT UKK DI EV ESUK UKK DI EV UKK DI	40         160           THE EKI ERYIKO         99           THE OF SK. EDNILKO         99           DOT SK. EDNILKO         90           DOT SK. EDNILKO         100           DOT SK. EDNIKE         100           DOT SK. EDNEKE         100           DOT SK. EDREKE         121           TGS RO. DOELKE         122           SCS RO. CPELKE         122           SCG LK. IPELLE         122           SCT RI. QEMERKE         122           SCT RI. QEMARARG         122           SCR RK. EDRARARG         122           SC
CSP-ecl CSP-sgl CSP-sg2 CSP-sg2 CSP-sg3 CSP-sg3 Pl0 OS-D4 OS-D5 OS-D1 OS-D2 Al0 HvirCSP2 SAP4 PEBmeIII SAP-1 ASP3C HmorCSP1 OlfP CSD-Harm CSP-HbraA1 CSP-NbraA2 CSP-NbraA3 CSP-NbraA4 CSP-NbraA4	* 100 * 120 GLADDAE LKAR DOTENE ALSEKTA VETT VF RPCPDCLAA DOTENE ALSEKTA VETT VF ANTVDCLKA PDT SNEALSEKTA VETT VF ANTVDCLKA PDT SNEALSEKTA VETT VF ANTVDCLKA PDT SNEALSEKTA VETT VF ANTADELSVIPD SNEALSEKTETELSK SNEAD SVEDT SNEALSEKTELSK SNEAD SVEDT SNEAD SVEDT SNEALSEKTELSK SNEAD SVEDT SNEAD SVEDT SNEAD SVEDT SNEAD SVEDT SNEAD SVEDT SNEAD SVEDT SNEAD SVEDT SVEDT SNEAD SVEDT SNEAD SVEDT SNEAD SVEDT SVEDT SNEAD SVEDT SNEAD SVEDT SVEDT SVEDT SNEAD SVEDT SNEAD SVEDT SVEDT SVEDT SNEAD SVEDT SVEDT SVEDT SVEDT SVEDT SVEDT SVEDT SVEDT SNEAD SVEDT SVED	* 1 KNK KEY (SPSK) NHK PDI AQUKA NHK PDI AQUKA	40         160           THE EKI ERYIKO         99           THE OF SK. EDNILKO         99           OGT SK. EDNILKO         90           DGT SK. EDNILKO         100           DGT SK. EDNIKE         100           DGT SK. EDREKE         121           DGT SK. EDREKE         122           SGS RO DOELKR         122           SGS RO GPELKK         122           SGT RI QEMKEK         122           SGT RI QEMKEK         122           SGT RI QEMKEKK         122           SGT RI QEMKEKK         122           SGT RI QEMKEKK         122           SGT RI QEMKAKG         122           SGK RK EDRARAKG         122           SKE RK EDRARAKG         122           SKE RK EDRARAKG
CSP-ec1 CSP-sg1 CSP-sg2 CSP-sg2 CSP-sg3 pl0 OS-D4 OS-D5 OS-D1 OS-D2 A10 HvirCSP2 SAP4 PEBmeIIII SAP-1 ASP3c EmorCSP1 OlfP CSP-Hzea HvirCSP3 CSP-NbraA2 CSP-NbraA3 CSP-NbraA4 SAP2 SAP2 SAP2 SAP2 SAP2 SAP2 SAP3 SAP3 SAP3 SAP3 SAP3 SAP3 SAP3 SAP3 SAP3 SAP3 SAP3 SAP3 SAP4	*         100         *         120           :         GL ADDAE         KAR DOT ENE ALSEKTA VETT VP         REPORT         VETT VP           :         GL ADDAE         KAR DOT ENE ALSEKTA VETT VP         REPORT         VETT VP           :         GL ADDAE         KAR DOT ENE ALSEKTA VETT VP         REPORT         KAR DOT ENE ALSEKTA VETT VP           :         AN TVD         KAR DOT ENE ALSEKTA VETT VP         REPORT         KAR DOT ENE ALSEKTA VETT VP           :         AN TVD         KAR DOT ENE ALSENTALSETE VELSK         KR         STAT         STAT           :         SNEAD         SNE ALSENTALSETE VELSK         KR         SNE ALSENTALSETE VELSK         SNE           :         SNE TDOT         SNE ALSENTALSETE VELSK         SNE         SNE         SNE           :         SNE TDOT         SNE ALSENTALSETE VELSK         SNE         SNE         SNE           :         SNE TDOT         SNE ALSENTALSETE VELSK         SNE         SNE         SNE         SNE           :         SNE TDOT         SNE ALSENTALSETE VELSK         SNE         SNE         SNE         SNE         SNE           :         SNE         SNE         SNE         SNE         SNE         SNE         SNE         S	* 1 C KNK EY ESPK KNK EY ESPK NHK PDI AQUKA U NHK PDI	40         160           THE DENI ERVIEQ
CSP-ec1 CSP-sg1 CSP-sg2 CSP-sg3 CSP-sg3 CSP-sg3 pl0 OS-D4 OS-D4 OS-D3 OS-D3 OS-D3 OS-D3 OS-D3 OS-D3 OS-D3 OS-D3 OS-D4 OS-D3 OS-D4 OS-D3 OS-D4 OS-D4 OS-D3 OS-D4 OS-D4 OS-D3 OS-D4 OS-D4 OS-D4 OS-D4 OS-D4 OS-D4 OS-D5 OS-D3 OS-D4 OS-D4 OS-D4 OS-D4 OS-D5 OS-D5 OS-D3 OS-D4 OS-D4 OS-D4 OS-D4 OS-D5 OS-D5 OS-D3 OS-D4 OS-D4 OS-D5 OS-D4 OS-D4 OS-D5 OS-D5 OS-D5 OS-D5 OS-D5 OS-D5 OS-D4 OS-D5 OS-D5 OS-D4 OS-D5 OS-D5 OS-D5 OS-D5 OS-D5 OS-D5 OS-D5 OS-D5 OS-D5 OS-D1 OS-D5 OS-D5 OS-D5 OS-D1 OS-D5	* 100 * 120 GLADDAE LKAR DD PENE ALSEKTA VETT VF RECEPCIENT KAR DD PENE ALSEKTA VETT VF AN TVD SLEKAPD I ENE ALSEKTA VETT VF AN TAD SLEKAPD I ENE ALSEKTA VETT VF AN TAD SLEKAPD I ENE ALSEKTA VETT VF SNEAD SLEXT DI ENE ALSEKTE TEL KH SNEAD SLEXT DI TINE ALSEKTE ALSE TEL AS TPD SLEAN PD TINE ALSEKTA ALSE KE AS TPD SLEAN PD TINE ALSEKTA ALSE KE SNEAD SLEXT DI TINE ALSEKTA ALSEKTE SNEAD SLEXT DI TINE ALSEKTA ALSEKTE SNEAD SLEXT DI TINE ALSEKTAR SLEXT SNEAD SLEXT DI TINE ALSEKTAR SLEXT SNEAD SLEXT DI TINE ALSEKTAR SLEXT SNEAD SLEXT SLEXT SLEXT SLEXT SLEXT SLEXT SLEXT SNEAD SLEXT SLEXT SLEXT SLEXT SLEXT SLEXT SLEXT SLEXT SLEXT SNEAD SLEXT	* 14 KNKEPI ESPK NHKPDI AQUKA NHKPDI AQUKA NHKPDI AQUKA INHKPDV AQUKA INHKP	40         160           THE EKI ERYIKQ
CSP-ec1 CSP-sg1 CSP-sg2 CSP-sg2 CSP-sg3 CSP-sg3 CSP-sg3 OS-D4 OS-D4 OS-D5 OS-D1 OS-D3 OS-D2 A10 HvirCSP2 SAP4 PEBmeIII SAP-1 ASP3C HmorCSP1 OlfP CSP-HbraA1 CSP-MbraA2 CSP-MbraA3 CSP-MbraA3 CSP-MbraA4 SAP2 CSP-MbraA3 CSP-MbraA4 SAP2 CSP-MbraA5 CSP-MbraA4 SAP2 CSP-MbraA5 CSP-MbraA6 CSP-MbraA1 CSP-MbraA6 CSP-MbraA6 CSP-MbraB1 CSP-MbraB1	* 100 * 120 GLADDAE LKAR DD PLENE ALSEKNEA VETT VF RESPONDELLEN ALDE ALSEKNEA VETT VF AND VD CLEAR ALDE ALSE ALSEKNEA VETT VF AND AD CLEAR ALDE ALSE ALSEKNEA VETT VF SN GAD CLEAR ALDE SNE ALSEKNE TELVEN SN GAD CLEAR ALDE THE ALSEN ALSE SNE SNE ALSE AS CTED CLEAR ALDE THE ALSEN SNE ALSE SNE SNE SNE SNE SNE SNE SNE SNE SNE S	* 14	40         160           THE EKI ERYIKO         99           THE OF DNLLKO         99           OCT SK EDNLKO         90           DOT SK EDNLKO         100           DOT SK EDNLKO         100           DOT SK EDNLKO         100           DOT SK EDNEKE         100           DOT SK EDNEKE         100           DOT SK EDREKE         100           DOT SK EDREKE         100           DOT SK EDREKE         100           DOT SK EDREKE         121           FGS RO DOELKE         122           FGS RO DOELKE         122           FGS RO DOELKE         122           SCS RO GELEKE         122           SCF RI OERLEN         122
CSD-ecl CSP-sgl CSP-sg2 CSP-sg2 CSP-sg3 CSP-sg3 CSD-ag3 pl0 OS-D4 OS-D5 OS-D1 OS-D2 OS-D2 Al0 HvirCSP2 SAP4 PEBmeIII SAP-1 ASP3C HmorCSP1 OlfP CSD-Harm CSP-MbraA1 CSP-MbraA2 CSP-MbraA3 CSP-MbraA3 CSP-MbraA4 SAP2 CSP-MbraB1 CSD-MbraB3 CSP-MbraB3 CSP-MbraB3	<ul> <li>100</li> <li>GLADDAE LKAR DOT ENE ALSEKTA VETT VF RP CPO CLOAD DO ENE ALSEKTA VETT VF AN TVD CLEKAR PD SNE ALSEKTA VETT VF AN TVD CLEKAR PD SNE ALSEKTA VETT VF AN TVD CLEKAR PD SNE ALSEKTA VETT VF SNEAD SLEAT DO ENE ALSEKTA VETT VF SNEAD SLEAT DO ENE ALSEKTE TELLEN SNEAD SLEAT DO ENE ALSEKTE ALSE TELLE SNEAD SLEAT DO ENE ALSEKTE ALSE TELLE SNEAD SLEAT DO ENE ALSEKTEN ALSE TELLE SNEAD SLEAT SLEAT SLEAT SLEAT SLEAT SNEAD SLEAT SLEAT SLEAT SLEAT SLEAT SLEAT SLEAT SNEAD SLEAT SLEAT SLEAT SLEAT SLEAT SLEAT SLEAT SNEAD SLEAT SLEAT</li></ul>	* 1. KNK KEY ESPKK NHKPDI AQUKA NHKPDI AQUKA NHKPDI AQUKA NHKPDI AQUKA NHKPDI AQUKA NHKPDI AQUKA NHKPDI AQUKA NHKPDI AQUKA SUB NHKPDI AQUKA SUB NHKPDI AQUKA SUB NHKPDI SUB SUB SUB SUB SUB SUB SUB SUB	40         160           THE EKI ERYIKO         99           THE OF INLEQ         99           OCT SK. EDNILKO         90           DOT SK. EDNILKO         100           DGT SK. EDNILKO         100           DGT SK. EDNILKO         100           DGT SK. EDNILKO         100           DGT SK. EDNIKE         100           DGT SK. EDNIKE         100           DGT SK. EDREE         121           TGS RO. DOELKR         122           SGS RO. GPELKK         101           SGT LK. IPELLSA         122           SGT K. IPELLSA         121           SGT K. IPELLSA         122           SGT K. IPELKK         122           SGT K. EDRAKAKG         122           SGT K. EDRAKAKG         122           SGT K. EDRAKAKG         122           SGT K. EDRAKAKG
CSP-ec1 CSP-sg1 CSP-sg2 CSP-sg2 CSP-sg3 pl0 OS-D4 OS-D5 OS-D1 OS-D2 A10 HvirCSP2 BAP4 PEBmeIIII SAP-1 ASP3c EmorCSP1 OIfP CSP-Hzea HvirCSP3 CSP-NbraA2 CSP-NbraA3 CSP-NbraA4 SAP2 CSP-NbraB3 CSP-NbraB1 CSP-NbraB3 CSP-NbraB2 CSP-NbraB2 CSP-NbraB2 CSP-NbraB3 CSP-NbraB3 CSP-NbraB3 CSP-NbraB3 CSP-NbraB3 CSP-NbraB4 SAP2 CSP-NbraB3 CSP-NbraB3 CSP-NbraB3 CSP-NbraB3 CSP-NbraB3 CSP-NbraB3 CSP-NbraB3 CSP-NbraB3 CSP-NbraB3 CSP-NbraB4 CSP-NbraB4 CSP-NbraB4 CSP-NbraB4 CSP-NbraB4 CSP-NbraB3 CSP-NbraB4 CS	*         100         *         120           GLADDAE         KAR DDT ENE ALSEKTA VETT VF         FRE PPO CLADA DDT ENE ALSEKTA VETT VF           AN TVD CLARA PDT ENE ALSEKTA VETT VF         AN TVD CLARA PDT ENE ALSEKTA VETT VF           AN TVD CLARA PDT ENE ALSEKTA VETT VF           AN TVD CLARA PDT ENE ALSEKTA VETT VF           AN TVD CLARA PDT ENE ALSEKTA VETT VF           AN TVD CLARA PDT ENE ALSEKTA VETT VF           AN TVD CLARA PDT ENE ALSEKTA VETT VF           SNEAD         SVI PDT ENE ALSEKTE TELVEN           SNEAD         SVI PDT ENE ALSEKTE TELVEN           SNEAD         SVI PDT ENE ALSEKTE TELVEN           AN TDD CLARA PDT TINE ACONENTAL AS URF           AS TPD CLARA PDT TINE ACONENTAL AS URF           GROTPD ALS PT PDT PDT ENE ACONENTAL AS URF           GROTPD ALS PT PDT PDT ENE ACONENTAL AS URF           GROTPD ALS PT PDT PDT ENE ACONENTAL AS URF           GROTPD ALS PT PDT PDT ENE ACONENTAL AS URF           GROTPD ALS PT PDT PDT ENE ACONENTAL AS UNF           GROTPD ALS PT PDT PDT ENE ACONENTAL AS UNF           GROTPD ALS PT PDT AND ACONENTAL AS UNF           GROTPD ALS PT PDT PDT ENE ACONE	* 1 C KNK KEY E SPKK V 7 C NHK PDT AQU KA V 7 C NK KEV E SO K K K K PDL E PU E K 7 C K K K PDL E PU E K 7 C K K K PDL E PU E K 7 C K K K PDL E PU E K 7 C K K K PDL E PU E K 7 C K K K PDL E PU E K 7 C K K K PDL E PU E K 7 C K K K PDL E PU E K 7 C K K K PDL E PU E K 7 C K K K PDL E PU E K 7 C K K K PDL E PU E K 7 C K K K PDL E PU E K 7 C K K K PDL E PU E K 7 C K K K PDL E FU E K 7 C K K K PDL E FU E K 7 C K K K PDL E FU E K 7 C K K K PDL E FU E K 7 C K K K PDL E FU E A 7 C K K K PD FU E K 7 C K K FU	40         160           THE EKI ERYIKQ
CSP-ec1 CSP-sg1 CSP-sg2 CSP-sg2 CSP-sg3 pl0 OS-D4 OS-D4 OS-D3 OS-D3 OS-D3 OS-D3 OS-D3 OS-D3 OS-D3 OS-D3 OS-D3 OS-D3 OS-D3 OS-D3 OS-D4 OS-D3 OS-D4 OS-D3 OS-D4 OS-D3 OS-D4 OS-D4 OS-D4 OS-D4 OS-D4 OS-D4 OS-D3 OS-D4 OS-D4 OS-D4 OS-D5 OS-D3 OS-D4 OS-D4 OS-D4 OS-D4 OS-D5 OS-D3 OS-D4 OS-D4 OS-D4 OS-D5 O	*       100       *       120         *       100       *       120         *       REPERDOR LIKAR DOLTENE ALSEKTRA VETT VE         *       REPERDOR LIKAR DOLTENE ALSEKTRA VETT VE         *       ANDADELIKAR POLTENE ALSEKTRA VETT VE         *       SNEADELIKAR POLTENE ALSEKTRE TELIKE         *       SNEADELIKAR POLTENE ALSEKTRE ALSEKTRE         *       SNEADELIKAR POLTENE ALSEKTRE         *       SNEA	* 14 C KNK KEY (S S S K K ) 7 C NHK PDT AQU KA ) 7 C NKK PDT S C C C C C C C C C C C C C C C C C C	40         160           THE EKI ERYIKQ
CSP-ec1 CSP-sg1 CSP-sg2 CSP-sg2 CSP-sg3 pl0 OS-D4 OS-D4 OS-D5 OS-D1 OS-D3 OS-D2 A10 HvirCSP2 SAP4 PEBmeIII SAP-1 ASP3c EmorCSP1 OlfP CSP-Hzea HvirCSP3 CSP-MbraA1 CSP-MbraA2 CSP-MbraA3 CSP-MbraA3 CSP-MbraA3 CSP-MbraA4 CSP-MbraA3 CSP-MbraA3 CSP-MbraA3 CSP-MbraA4 CSP-MbraA3 CSP-MbraA3 CSP-MbraA4 CSP-MbraA4 CSP-MbraB1 CSP-MbraB1 CSP-MbraB2 CSP-MbraB2 CSP-MbraB3 CSP-MbraB3 CSP-MbraB4 HvirCSP1 SAP5	*         100         *         120           GLADDAE         EKAR DD. ENE ALSEKTA VETT VF           RPCPDOCLEAN DD.         ENE ALSEKTA VETT VF           AN TVD.SL.KAPDD.         ENE ALSEKTAP.           SN TAD.SL.KAPDD.         ENE ALSEKTAP.           SN TAD.SL.KAPDD.         ENE ALSEKTAP.           SN TAD.SL.KAPD.         ENE ALSEKTAP.           SN TAD.SL.KAPD.         ENE ALSEKTAP.           SN TAD.SL.KAPD.         ENE ALSEKTAP.           SN TPD.SL.KAPD.         ENE ALSEKTAP.           GK TPD.SL.KAPD.         ENE ALSEKTAP.           AS TPD.SL.KAPD.         TINE ACONENTAL AS VIET           AS TPD.SL.KAPD.         TINE ACONENTAL AS VIET           AS TPD.SL.KAPD.         TINE ACONENTAL AS VIET           GR TPD.SL.KAPD.         TINE ACONENTAL AS VIET           GR TPD.SL.KAPD.         TINE ACONENTAL AS VIET           GR TPD.SL.KAPD.         PD.TINE ACONENTAL AS VIET	* 14 C KNK KPE ( ESPK ( ) 77 C NHK PDT ( AUKA) ( ) 7 C NKK PDT ( ) 7 C NKK PT ( )	40         160           THE EKI ERYIKO         99           THE OF SK. EDNILKO         99           DOT SK. EDNILKO         100           DOT SK. EDRIKE         100           DOT SK. EDRIKE         100           DOT SK. EDRIKE         121           FOR RO. DOELKE         122           FOR RO. DOELKE         122           SCS RO. CPELKE         122           SCR RE. ERCAQUKE         122           NUT OF KENERARANG         122           NUT OF KENERARANG         122           SCR
CSD-ecl CSP-eg2 CSP-sg2 CSP-sg2 CSP-sg3 CSD-ag3 pl0 OS-D4 OS-D5 OS-D1 OS-D2 Al0 HvirCSP2 SAP4 PEBmeIII SAP1 ASP3C HmorCSP1 OlfP CSD-Harm CSP-MbraA1 CSP-MbraA2 CSP-MbraA3 CSP-MbraA3 CSP-MbraA4 SAP2 CSP-MbraB1 CSD-MbraB1 CSD-MbraB3 CSP-MbraB3 CSP-MbraB3 CSP-MbraB3 CSP-MbraB3 CSP-MbraB4 HvirCSP1 SAP3 CSP-MbraB3 CSP-MbraB4 CSP-MbraB3 CSP-MbraB3 CSP-MbraB4 CSP-MbraB3 CSP-MbraB4 CSP-MbraB3 CSP-MbraB4 CSP-MbraB4 CSP-MbraB3 CSP-MbraB4 CSP-MbrA4 CSP-MbraB4	*       100       *       120         :       GLADDAE       KAR PD. TENE AR SERVAR VETT VF         :       R.P. CPO.       CARA PD. TENE AR SERVAR VETT VF         :       AN TVD       KA PD. ENE AR SERVAR VETT VF         :       AN TVD       KA PD. ENE AR SERVAR VETT VF         :       AN TVD       KA PD. ENE AR SERVAR VETT VF         :       AN TVD       KA PD. ENE AR SERVAR VETT VF         :       SN TAD	* 1 KNK KEY GEPER NHKPDI AQUKA NHKPDI AQUKA NHKPDI AQUKA NHKPDI AQUKA NHKPDI AQUKA NHKPDI AQUKA KEYDL EPUER KEKPLI EPUER KEKPLI EPUER KEKPLI EPUER KEKPLI EPUER KEKPLI EPUER KEKPLI EPUER KEKPLI EPUER KEKPLI EPUER NEROL KEISA NEROL KEISA NEROL KEISA NEROL KEISA NEROL KEISA NEROL KEISA NEROL KEISA KELDI EPUER KNELDI EPIER KNELDI EPIER KNELD	40         160           THE EKI ERYIKO         99           THE OFFICE         99           OFFICE         910           DOT SK. EDKEKE         100           DOT SK. EDKEKE         100           DOT SK. EDKEKE         100           DOT SK. EDKEKE         100           DOT SK. EDREKE         121           TGS RO. DOELKE         122           SCS RO. CPELKE         122           SCG K. UPELLSA         122           SCT RI. QEMERKE         122           SCT RI. QEMERANG         122           SCT RI. QERARAG         122           SCR RK. EDRARAKG         122           SCG RK. EDRARAAG
CSP-ec1 CSP-sg1 CSP-sg2 CSP-sg2 CSP-sg3 CSP-sg3 pl0 OS-D4 OS-D5 OS-D1 OS-D2 A10 HvirCSF2 SAP4 PEBmeIIII SAP-1 ASF3c BmorCSF1 OlfP CSP-HbraA1 CSP-MbraA2 CSP-MbraA3 CSP-MbraA3 CSP-MbraA3 CSP-MbraA3 CSP-MbraA4 SAF2 CSP-MbraB1 CSP-MbraB3 CSP-MbraB3 CSP-MbraB3 CSP-MbraB3 CSP-MbraB3 CSP-MbraB4 SAF3 SAF3 SAF3 SAF3 SAF3 SAF5 CSP-c2 BmorCSF2 CSP-c2 BmorCSF2 CSP-c2 CSP	*         100         *         120           GL ADDAE         KAR PD. TENE ALSEKTA VETT VF         F	*       1         KNKEPEY ESPEK       7         INHKPDI AQUKA       7         INKENDI AQUKA       7         INKENDI FUNCA       7         INKROL FUNCA       7         INKELDI FEIAA       7         INKELDI FEIAA       7         INKEI FEIAA       7         INKEI FEIAA       7         INKEI FEIAA       7         INKEI FEIAA       7         INHEDAY KEIA <td< td=""><td>40         160           THE DENI ERVINCO         99           THE OFF SK. EDNERE         100           DOT SK. EDNERE         100           DOT SK. EDNERE         100           DOT SK. EDRERE         121           TGS RO. DOELKR         122           SGS RO. DOELKR         122           SGS RO. DOELKR         122           SGS RO. GERLER&lt;</td>         122           SGT LK. UPELLSA         122           SGT LK. UPELLSA         122           SER LE RC-QACKE         122           DOK KDKIERIK         122           DOK KDKIERIK         122           DOK KDKIERIK         122           DOK KDRIBER         122           DOK KERV EE-SAKKL         122           DOK RK EDRARANG         122           KGD RK EDRARANG         121           DOK RK EDRARANG         100           TGN RK EDRARA</td<>	40         160           THE DENI ERVINCO         99           THE OFF SK. EDNERE         100           DOT SK. EDNERE         100           DOT SK. EDNERE         100           DOT SK. EDRERE         121           TGS RO. DOELKR         122           SGS RO. DOELKR         122           SGS RO. DOELKR         122           SGS RO. GERLER<
CSD-ec1 CSP-eg2 CSP-sg1 CSP-sg2 CSP-sg2 CSP-sg3 pl0 OS-D4 OS-D4 OS-D5 OS-D1 OS-D3 OS-D2 A10 HvirCSP2 SAP4 PEBmeIII SAP-1 ASP3c EmorCSP1 OIfP CSD-Harm CSD-H2ea HvirCSP3 CSP-MbraA1 CSP-MbraA2 CSP-MbraA3 CSP-NbraA3 CSP-NbraA3 CSP-NbraA3 CSP-NbraB3 CSP-NbraB1 CSP-NbraB1 CSP-NbraB1 CSP-NbraB3 CSP-NbraB3 CSP-NbraB3 CSP-NbraB3 SAP5 CSP-sc2 SAP1 SAP5	*         100         *         120           GL ADDAE         KAR DD TENE ALSEKTA VETT VF         FRE PPO CLEAR DD TENE ALSEKTA VETT VF           AN TVD CLEAR PD TENE ALSEKTA VETT VF         AN TVD CLEAR PD TENE ALSEKTA VETT VF           AN TVD CLEAR PD TENE ALSEKTA VETT VF           AN TVD CLEAR PD TENE ALSEKTA VETT VF           AN TVD CLEAR PD TENE ALSEKTA VETT VF           AN TAD         KALPD TENE ALSEKTA VETT VF           AN TAD         KALPD TENE ALSEKTA VETT VF           SN TAD         SVIPD TENE ALSEKTE TELSKH           SN TAD         SVIPD TENE ALSEKTE TELSKH           GK TPD SL AALPD TINE ALSEKTE TELSKH           AS TPD SL AALPD TINE ALSEKTE TELSKH           AS TPD SL AALPD TINE ALSEKTA AS TRF           GR TPD ALSE TRF         ST ANTER AS TRF           GR TPD ALSE TFD DU TINE ALSEKTATALAS TRF	* 14 KNKEPI ESPK NHKPDI AQUKA NHKPDI AQUKA NHKPDI AQUKA NHKPDI AQUKA NHKPDI AQUKA NHKPDI AQUKA NHKPDI AQUKA NHKPDI ESUK KKPDI ESUK KKPDI ESUK NKKPI ESUK NKKI E	40         160           THE EKI ERYIKQ
CSP-ec1 CSP-sg1 CSP-sg2 CSP-sg2 CSP-sg3 CSP-sg3 pl0 OS-D4 OS-D4 OS-D3 OS-D3 OS-D3 OS-D3 OS-D3 OS-D3 OS-D3 OS-D3 OS-D3 OS-D3 OS-D3 OS-D4 OS-D3 OS-D4 OS-D3 OS-D4 OS-D3 OS-D4 OS-D4 OS-D4 OS-D3 OS-D4 OS-D4 OS-D3 OS-D4 OS-D4 OS-D4 OS-D3 OS-D4 OS-D3 OS-D3 OS-D3 OS-D4 OS-D3 OS-D4 OS-D3 OS-D4 OS-D3 OS-D4 OS-D3 OS-D4 OS-D4 OS-D3 OS-D4 OS-D3 OS-D4 OS-D3 OS-D4 OS-D3 OS-D4 OS-D3 OS-D4 OS-D3 OS-D4 OS-D3 OS-D4 OS-D4 OS-D4 OS-D4 OS-D4 OS-D4 OS-D4 OS-D4 OS-D4 OS-D4 OS-D4 OS-D4 OS-D5 OS-D4 OS-D3 OS-D4 OS-D4 OS-D3 OS-D4	*       100       *       120         *       100       *       120         *       100       *       120         *       100       *       120         *       RP CPD0       CLADAD       PD1       PNE AN CREAK VETT       VP         AN TVD CLADAD       KAUPD1       PNE AN CREAK VETT       VP         AN TVD CLADAD       KAUPD1       PNE AN CREAK VETT       VP         AN TVD CLADAD       KAUPD1       PNE AN CREAK VETT       VP         SN TAD       SVI PD1       PNE AN CREAK VETT       VF         SN TAD       SVI PD1       PNE AN CREAK VETT       VF         GK TPD2       SAN PD1 <tine an="" creak="" td="" vett<="">       PN         AS TPD3       SAN PD1<tine af="" an="" creak="" td="" vett<="">       PN         AS TPD3       SAN PD1<tine af="" an="" creak="" td="" vett<="">       PN         AS TPD3       SAN PD1<tine af="" an="" creak="" td="" vett<="">       PN         AS TPD3       SAN PD1<tine af="" an="" creak="" td="" vett<="">       PN         AS TPD3       SAN PD1<tine af="" an="" creak="" td="" vett<="">       PN         GK TPD3       SAN PD1<tine af="" an="" creak="" td="" vett<="">       PN         GK TPD3       SAN PD1<tine af="" an="" creak="" td="" vett<="">       PN         GK TPD3       SAN PD1<tine a<="" td=""><td>* 14 C KNK KEY (599 KK ) 77 C NHK PDT AQU KA (50 C NHK PDT (50 C K K PDL (50 C K K PDC (50 C</td><td>40         160           THE EKI ERYIKQ</td></tine></tine></tine></tine></tine></tine></tine></tine></tine>	* 14 C KNK KEY (599 KK ) 77 C NHK PDT AQU KA (50 C NHK PDT (50 C K K PDL (50 C K K PDC (50 C	40         160           THE EKI ERYIKQ
CSP-ec1 CSP-sg1 CSP-sg2 CSP-sg3 CSP-sg3 pl0 08-D4 08-D5 08-D1 08-D3 08-D2 A10 HvirCSP2 SAP4 PEBmeIII SAP-1 ASP3C EmorCSP1 OlfP CSP-H2ea HvirCSP3 CSP-MbraA1 CSP-MbraA2 CSP-MbraA3 CSP-MbraA3 CSP-MbraA3 CSP-MbraA3 CSP-MbraA3 CSP-MbraA3 CSP-MbraA3 CSP-MbraA3 CSP-MbraA3 CSP-MbraA3 CSP-MbraA3 CSP-MbraA3 CSP-MbraA3 CSP-MbraA3 CSP-MbraA3 CSP-MbraA3 CSP-MbraA3 CSP-MbraB2 CSP-MbraB1 CSP-MbraB1 CSP-MbraB1 SAP3 SAP5 CSP-ec2 BmorCSP2 SAP1 SAP	$\begin{array}{c c c c c c c c c c c c c c c c c c c $	*       1         V.KNKEPE BESKK       25         V.KNKEPE BESKK       25         V.NKKEPI BESKK       25         V.NKKEPI ADIKA       25         V.NKKEPI ADIKA       25         V.NKKEPI ADIKA       26         V.NKKEPI ADIKA       26         V.NKKEPI ADIKA       27         V.NKKEPI ADIKA       26         V.KKEPI EDIK       26         V.KKEDI EDIK       26         V.KKEDI EDIKA       26         V.KKEDI EDIKA       27         V.KKEDI EDIKA       26         V.KKEDI EDIKA       27         V.KKEDI EDIKA       26         V.KKEDI EDIKA       27         V.KKEDI EDIKA       27         V.KKEDI EDIKA       27         V.KKEDI EDIKA       27         V.KKEDI EKEROLOKAN       27         V.KKEDI EKEROLOKAN       27         V.KKEDI RETA       27	40         160           THE EKI ERYIKQ

Fig. 1 Multiple alignment showing the absolutely conserved domain in insect CSPs: Cx(6,8)Cx(18)Cx(2)Cx(3). All acronyms in the figure refer to the protein names listed in Table 1.

Peptide ID	Previously identified insect CSPs						
agCP10968	SAP-1	ASP3c	SAP2	CSP-Harm	HvirCSP2	OS-D3	CSP-sg1
	(6e-18)	(7e-18)	(6e-17)	(1e-16)	(2e-16)	(2e-16)	(2e-16)
	OS-D1	CSP-sg4	CSP-sg2	$\operatorname{CSP-MbraA6}$	OS-D4	CSP-Hzea	OS-D5
	(2e-16)	(3e-16)	(3e-16)	(3e-16)	(4e-16)	(5e-16)	(3e-16)
	SAP4	PEBmeIII	CSP-sg5	OS-D2	A10	CSP-ec3	CSP-MbraA3
	(7e-16)	(9e-16)	(9e-16)	(2e-15)	(2e-15)	(2e-15)	(2e-15)
	CSP-MbraA1	CSP-MbraA2	CSP-sg3	p10	CSP-MbraA5	SAP3	
	(2e-15)	(3e-15)	(3e-15)	(3e-15)	(7e-15)	(8e-15)	
agCP11079	SAP-1	PEBmeIII	ASP3c	p10	HvirCSP2	SAP4	A10
	(1e-45)	(1e-34)	(2e-34)	(4e-28)	(3e-27)	(1e-26)	(2e-26)
	OS-D2	$\operatorname{CSP-MbraA6}$	OS-D3	SAP3	BmorCSP1	SAP5	CSP-sg1
	(6e-26)	(2e-23)	(4e-23)	(7e-23)	(8e-23)	(9e-23)	(1e-22)
	CLP-1	CSP-MbraA3	CSP-sg4	HvirCSP1	CSP-sg2	OS-D4	OS-D5
	(2e-22)	(3e-22)	(4e-22)	(5e-22)	(6e-22)	(7e-22)	(7e-22)
	CSP-sg4	OS-D1	$\operatorname{CSP-MbraA2}$	CSP-sg3	SAP2	$\operatorname{CSP-MbraA3}$	agCP11435
	(1e-21)	(1e-21)	(2e-21)	(3e-21)	(4e-21)	(4e-21)	(5e-08)
agCP11481	ASP3c	PEBmeIII	A10	SAP4	HvirCSP2	SAP5	SAP3
	(4e-28)	(3e-27)	(3e-26)	(3e-25)	(4e-25)	(4e-24)	(2e-23)
	CSP-sg2	SAP-1	CSP-sg1	OS-D2	$\operatorname{CSP-MbraA6}$	CLP-1	CSP-sg5
	(1e-22)	(1e-22)	(2e-22)	(3e-22)	(7e-22)	(1e-21)	(2e-21)
	CSP-sg4	$\operatorname{CSP-MbraA3}$	CSP-sg3	$\operatorname{CSP-MbraA2}$	OS-D3	HvirCSP1	$\operatorname{CSP-MbraA1}$
	(2e-21)	(9e-21)	(1e-20)	(2e-20)	(3e-20)	(4e-20)	(6e-20)
	$\operatorname{CSP-MbraA3}$	$\operatorname{CSP-MbraB1}$	SAP2	$\operatorname{CSP-MbraA4}$	$\operatorname{CSP-MbraA5}$	BmorCSP1	agCP11435
	(7e-20)	(8e-20)	(1e-19)	(2e-19)	(2e-19)	(3e-19)	(1e-06)
agCP11484	SAP-1	ASP3c	PEBmeIII	HvirCSP2	p10	$\operatorname{CSP-MbraA6}$	OS-D2
	(3e-55)	(1e-31)	(5e-32)	(6e-29)	(4e-28)	(3e-27)	(9e-27)
	SAP4	$\operatorname{BmorCSP1}$	CLP-1	SAP2	A10	SAP3	SAP5
	(1e-26)	(2e-26)	(2e-26)	(1e-25)	(6e-25)	(4e-24)	(6e-24)
	CSP-Hzea	$\operatorname{CSP-MbraA2}$	HvirCSP1	$\operatorname{CSP-MbraA1}$	$\operatorname{CSP-MbraA3}$	CSP-Harm	$\operatorname{CSP-MbraA4}$
	(3e-23)	(5e-23)	(5e-23)	(7e-23)	(7e-23)	(9e-23)	(2e-22)
	$\operatorname{CSP-MbraA5}$	CSP-sg1	CSP-sg4	CSP-sg2	OS-D3	OS-D1	CSP-ec1
	(2e-22)	(4e-22)	(8e-22)	(1e-21)	(2e-21)	(3e-21)	(3e-21)
	CSP-sg5	agCP11435					
	(3e-21)	(2e-08)					
agCP11532	RH70879	ASP3c	SAP5	CSP-sg4	$\operatorname{CSP-MbraA6}$	CSP-sg5	CSP-sg2
	(5e-30)	(4e-10)	(6e-08)	(6e-08)	(1e-07)	(1e-07)	(2e-07)
	HvirCSP2	CSP-sg1	SAP2	CSP-ec3	$\operatorname{BmorCSP2}$	CSP-sg3	OS-D3
	(2e-07)	(3e-07)	(4e-07)	(7e-07)	(7e-07)	(9e-07)	(1e-06)
	CLP-1	$\operatorname{CSP-MbraA3}$	$\operatorname{CSP-MbraA1}$	$\operatorname{CSP-MbraA2}$	$\operatorname{CSP-MbraA4}$	SAP4	$\operatorname{CSP-MbraA5}$
	(1e-06)	(1e-06)	(1e-06)	(1e-06)	(1e-06)	(2e-06)	(2e-06)
	OS-D2	OS-D1	A10	$\operatorname{CSP-Hzea}$	HvirCSP3	CSP-ec1	
	(3e-06)	(5e-06)	(1e-05)	(1e-05)	(1e-05)	(3e-05)	
agCP11545	SAP-1	PEBmeIII	ASP3c	p10	HvirCSP2	A10	OS-D2
	(3e-40)	(2e-33)	(2e-32)	(1e-30)	(6e-30)	(1e-28)	(1e-27)
	SAP4	$\operatorname{CSP-MbraA6}$	CSP-sg1	BmorCSP1	SAP3	CSP-sg4	CSP-sg2
	(3e-26)	(1e-25)	(3e-24)	(3e-24)	(3e-24)	(6e-24)	(6e-24)

Table 2 Anopheles CSP Candidates Found by CSPMOT and BLAST (E\*-values<0.0001)

Dentide ID			Duorriouslas	identified inco.	at CCDa				
	r reviously identified insect C5r's								
agCP11545	OS-D3	SAP5	CSP-sg5	CSP-sg3	HvirCSP1	OS-D4	OS-D1		
	(1e-23)	(2e-23)	(3e-23)	(3e-23)	(4e-23)	(5e-23)	(7e-23)		
	$\operatorname{CSP-MbraA3}$	$\operatorname{CSP-MbraA5}$	CLP-1	CSP-ec1	CSP-MbraA2	$\operatorname{CSP-Hzea}$	agCP11435		
	(2e-22)	(2e-22)	(2e-22)	(5e-22)	(7e-22)	(1e-21)	(7e-09)		
agCP6514	no matches <sup>a</sup>								
agCP12965	no matches <sup>a</sup>								
$agCP11435^{b}$ A10	OS-D1	OS-D3	OS-D5	OS-D4	SAP-1	SAP4			
	(2e-12)	(5e-10)	(7e-10)	(1e-09)	(2e-09)	(2e-08)	(3e-08)		
	PEBmeIII	OS-D2	HvirCSP2	ASP3c	CSP-sg5	CSP-ec2	CSP-sg4		
	(3e-08)	(4e-08)	(8e-08)	(1e-07)	(3e-07)	(5e-07)	(5e-07)		
	CSP-sg3	CSP-sg2	CSP-sg1	CSP-ec1	CSP-ec3	p10	$\operatorname{CSP-MbraA6}$		
	(6e-07)	(9e-07)	(9e-07)	(2e-06)	(2e-06)	(5e-05)	(5e-05)		
	CSP-Hzea	CSP-Harm	CSP-MbraA3	CSP-MbraA1					
	(1e-04)	(1e-04)	(1e-04)	(1e-04)					

 Table 2 Continued

<sup>a</sup> No significant hits obtained by BLAST;

<sup>b</sup> BLAST identified;

\* The E value is a parameter that describes the number of hits one can "expect" to see just by chance when searching a database of a particular size. All acronyms of insect CSPs refer to protein names listed in Table 1.

Table 3 Chromosomal Location, New ORFs, Signal Peptides, and Biochemical Propertie
of Anopheles CSP Candidates

Celera_ID	GB_ID	Chrom	Scaffold	Original	New ORF	Signal	MW	pI	Hydrophobic
			No.	length (a.a.)	length (a.a.)	peptide	(kDa)		a.a. (%)
agCP10968	EAA12703	3R	AAAB01008964	127	109	none	12.3	9.5	25.7
agCP11079	EAA12353	3R	AAAB01008964	143	127	1 - 17	14.8	5.4	33.1
agCP11481	EAA12591	3R	AAAB01008964	137	123	1 - 19	14.3	9.4	29.3
agCP11484	EAA12322	3R	AAAB01008964	149	127	1 - 17	14.7	8.6	33.1
agCP11532	EAA12601	3R	AAAB01008964	150	117	1-33	12.9	9.8	41.0
agCP11545	EAA12338	3R	AAAB01008964	141	126	1 - 17	14.6	8.6	31.0
agCP11435	EAA12702	3R	AAAB01008964	102	137	1-16	15.7	5.0	35.0
agCP12965	EAA05664	3L	AAAB01008834	173	137	none	14.1	3.5	23.4
agCP6514	EAA10937	2L	AAAB01008960	132	117	1-19	13.0	8.6	25.6

Peptide ID	Predicted secondory structure (%)			cture (%)	Globularity
	Helix	Sheet	Loop	Class	-
agCP10968	25.70	15.50	58.80	mixed	appears not to be globular
agCP11079	72.40	0.00	27.60	all-alpha	may be globular, but it is not as compact as a domain
agCP11481	71.50	0.00	28.50	all-alpha	may be globular, but it is not as compact as a domain
agCP11484	71.70	1.60	26.80	all-alpha	appears as compact, as a globular domain
agCP11532	72.70	0.00	27.40	all-alpha	appears as compact, as a globular domain
agCP11545	68.20	2.40	29.40	all-alpha	may be globular, but it is not as compact as a domain
agCP11435	75.90	0.00	24.10	all-alpha	may be globular, but it is not as compact as a domain
agCP12965	13.10	16.80	70.10	mixed	appears as compact, as a globular domain
agCP6514	10.30	0.00	89.70	mixed	appears not to be globular

\* Prediction server—http://www.sbg.bio.ic.ac.uk/3dpssm.

#### Homology modelling

The NMR solution structure of chemosensory protein Csp2 of *Mamestra brassicae* (1K19\_A) was retrieved from Protein Data Bank (PDB ID: 1K19). The 3-D structure of this model molecule is shown in Figure 2A, which is characterized by two disulfide bonds (CysI-CysII, CysIII-CysIV), instead of three disulfide bonds (CysI-CysIII, CysIII-CysV), and CysIV-CysVI) characterized by OBPs. The model has a typical hydrophobic core that is supposed to act as a pocket for ligand binding. The pocket is formed by hydrophobic amino acids, surrounded by hydrophilic amino acids.

Homology modelling of the Anopheles CSP candidates was made in Swiss-PdbViewer (Figure 2). The figures showed that most candidates folded similarly as 1K19\_A. However, a structurally weak linkage occurred between the two disulfide bridges of agCP11435 (Figure 2D), though a hydrophobic pocket was formed in the core of its structure. Further sequence analysis was carried out, which showed that the structural abnormality might be caused by a surplus sequence insertion (GRLACLALVL; Figure 3).

## Discussion

Sequence similarity in CSP primary structure is significantly higher than that in OBPs. OBP sequences of the same species could be less similar than those of different species, whereas CSP sequences are more conserved at species level as well as between phylogenetically distant groups. In the present study, the sequence similarity between agCP11484 and other CSP candidates is: 77%, 76%, 47%, 39%, 21%, 15%, 13%, and 8%, respectively. Moreover, the positions of two disulfide bridges were highly conserved. The structural conservation should be extremely important in forming a strong hydrophobic core that function as an odorant-binding site.

Thirty-eight OBP candidates have been annotated in *D. melanogaster*, whereas only 29 OBP candidates have been conceptually identified in *A. gambiae*, based on sequencing and genome analysis (10, 24-26). In the present study, we located nine CSP candidates in *Anopheles* genome, among which five candidates, *i.e.* agCP11079, agCP11481, agCP11484, agCP11532, and agCP11545, are the most possible *Anopheles* CSPs, considering the sequence similarity in their primary structure, the biochemical properties such as hydrophobicity and molecular weight (<15 kDa), in particular the secondary structure and 3-D



Fig. 2 Homology modelling of Anopheles CSP candidates with 1K19\_A\_MbraCSP (PDB ID: 1K19) as the model. A. 3D-structure of 1K19\_A\_MbraCSP showing the position of two disulfide bridges; B-J: predicted 3D-structure of Anopheles CSPs, including agCP11079, agCP10968, agCP11435, agCP11481, agCP11484, agCP11532, agCP 11545, agCP12695, and agCP6514. The figures are generated in Swiss-PdbViewer.

I

CSP-ec1	:	EGL CAPDAEELKKA IPD AL ENE CAKCSEK OKAGVETTIVPL I KNKPE IWESP KKKYD PTHKYEK I YER-Y	:	96
CSP-ec3	:	NRP CP PDGQELKDA IPDALENE CAKCSEKQKAGVETTIVFLIKNKPEVWESPKKKYDPTHKYQTFYDN-L	:	96
CSP-sg1	:	EANCTVDGKELKKAVPDALSNECAKCNDKQKEGTKKVLKHLINHKPDIWAQLKAKYDPDGTYSKKYED-K	:	103
CSP-ag2	:	EANCTADGKELKKAVPDALSNECAKCNDKOKEGTKKVLKHLINHKPDIWAQLKAKYDPDGTYSKKYED-K	:	103
CSP-ag4	:	ESWCTADGKELKSVIPDALSNECAKCNEKOKEGTKKVLKHLINHKPDVWAQLKAKYDPDGTYSKKYED-R	:	103
CSP-sg5	:	ESNCTADGKELKKD IPD AL SNE CAKCNEK QKE GTKKVLKHL I NH KPD VWAQLKAKYD PD GTYSKKYED-R	:	103
CSP-sg3	:	DTNCTDDGKELKSVIPDALSNECAKCNEKQKEGTKKVLKHLINHKPDIWAQLKAKYDPDGTYSKKYED-R	:	103
p10	:	EGKCTPDGEELKSHVSDALQNDCAKCSDKORAGAEKVINFLYNKKKPMWESLQKKYDPENTYVTKYAD-R	:	124
os-D4	:	DASCTPDGKELKAA IPDAL THE CAQCHEKQKAGAEKV IRFLIKEKPDLWEPLEKKYDPTGSPRQKYDQ-E	:	119
OS-D5	:	DASCTPDGKELKAAIPDALTNECAQCNEKQKAGAEKVIRFLIKEKPDLWTPLEKKYDPTGSPRQKYDQ-E	:	119
CS-D1	:	DASCTPDGKELKAVIPDALTNECAKCNEKOKAGAEKVIKFLVKEKPDLWEPLEKKYDPSGSPRQKYGP-E	:	97
OS-D3	:	DTFCTADGKELKAAIPDALTNECAKCNEKOKNGAEKVIRFLIKEKPDLWTPLEKKYDPNGTYRQKYGE-E	:	119
OS-D2	:	DASCTPDGKELKVSIPDALVTDCSKCNEKQKEGSNKVIRPLIQKKEDLWKPLQAKYDPEGTYLKKHPE-L	:	117
A10	:	TGPCTPDAKMLKEILPDAIQTDCTKCTEKQRYGAEKVTRHLIDNRPTDWERLEKIYDPEGTYRIKYQE-M	:	147
agCP11435	:	VGPCTPDGRELKGRLACLALVLDHNLPDALMSDCVKCSEKORIGSDKVIKFIVANRPDDPAILEQLYDPTGEYRRKYMQ-S	:	131
agCP11484	:	QGRCTPDGNELKRILPDALQTNCEKCSEKQRDGAIKVINYLIQNRKDQWDVLQKKFDPENKYLEKYRG-Q	:	118
SAP-1	:	QGRCTPDGNELKRILPDALQTNCEKCSEKQRDGAIKVINYLIQNRKDQWDVLQKKPDPENKYLEKYRG-Q	:	118
agCP11079	:	EGRCTPDGNELKKILPEALQTNCEKCSEKORSGAIKVINYVIENRKEQWDALQKKYDPENLYVEKYRE-E	:	118
agCP11545	:	TGRCTPDGNELKRILPDALKTDCAKCSEKQKSGTEKVINYLIDNRKDQWENLQKKYDPENIYVNKYRE-D	:	118
PEBmeIII	:	NGKCTPEGRELKKSLPDALKTECSKCSEKORONTOKVIRYIIENKARGVEAA QAKYDPEEIYIKRYRQPR	:	114
agCP11481	:	KGPCTQEGRELKKTLPDALKTWCEKCSEKQRTSSRKVIAHLEERKPQEWKKLLDKYDPEGIYKSKPEK	:	118
ASP3c	:	EGRCTAEGNELKRVLPDALATDCKKCTDKQREVIKKVIKPLWENKPELWDSLANKYDPDKKYRVKPEE-E	:	122
BmorC3P1	:	KGKCTPEGKELKDHLQEALETGCEKCTEAQEKGAETS ID YLIKNELE IWKELTAHFDPDGKWRKKYED-R	:	117
CLP-1	:	KGKCTPEGKELKEHLQEAIENGCEKCTEAQEKGAYTVIEHLIKNEIEIWRQLADKFDPERKYRKKYED-R	:	120
CSP-Harm	:	RGKCS PEGKELKEHLQDAIETGCSKCTEAQEKGAYKVIEHLIKNELDIWRELAAKYD PKGDWRKKYED-R	:	117
CSP-Hzea		RGKCS PEGKELKEHLODAIETGCS KCTEAOEKGAYKVIEHLIKNELDIWRELAAKYD PKGDWRKKYED-R	:	118
HvirCSP3	:	RGKCS PEGKELKEHLQDAIETGCS KCTEAQEKGAYKVIEHLIKNELDIWRELTAKYD PKGDWR-KYED-R	:	96
CSP-MoraA1	:	RGKCS PEGKELKEHLODAIENGCKKCAENOEKGAYRVIEHLIKNEIEIWRELTAKYD PTGNWRKKYED-R	:	101
CSP-MoraA2		RGKCSPEGKELKEHLODAIENGCKKCAENOEKGAYRVIEHLIENEIEIWRELTAKYDPTGNWRKKYED-R	:	101
CSP-MbraA5	:	RGKCS PEGKELKEHLODAIENGCKKCTENGEKGAYRAIEHLIKNEIEIWRELTAKYD PTGNWRKKYED-R	:	101
CSP-MoraA3		RGKCS PEGKELKEHLOD A I ENGCKKCTEN GEKGA YRV IE HLIKNE IE IWREL TA KYD PTGNWRKKYED-R	:	101
CSP-MoraA6		RGKCS PEGKELKEHLODAIENGCKKCTENOEKGAYRVIEHLIKNEIEIWRELTAKYD PTGNWRKKYED-R		117
CSP-MoraA4		RGKCSPEGKELKEHLODAIENGCKKCTENOEKGAYRVIEHLIKNEIEIWRELTAKYDPTYNWRKKYED-E		101
SAP2	-	RGKCTPEGKELKEHLRDAIETGCKKCTKPOEEGATKVIDFLIKNKLEVWRELVAKFDPEGKWRKKYED-R		117
CSP-MoraB1		OGK CAPDA KELKEH TREALENE COKCTETOKNOTREV I GHL I NHEDA YMKEL TA KYD POSKETA KYEK-E		1.02
CSP-MoraB3		OGKCAPDAKELKEHIREALENECGKCTETOKNGTREVIGHLINHEDAYWEELTAKYDPOSKFTAKYEK-E		102
CSP-MoraB2		OGKCAPDGKELKEHIKEALENE CGKCTDAOKKGTREVIAHLINHEED FWNELTAKFD PERKFTAKYEK-E		102
CSP-MoraB4	-	EGKCAPDGKELK		102
HyirCSP1	-	OCK CADDA KELK		108
SAD3		OGY CADDA VELV		120
SAPS	-	GORCTARGELK	;	120
HyirCSP2	2	DODOTED DOKALKFTLDD ALENE CSECTER OKA GSDEVID VLVNK DODLAKEL SA KYD DNEI YOD DYKD-K		110
SSP4		CORCEPTION AND A DESCRIPTION OF A DESCRIPTION OF A DESCRIPTION OF A DESCRIPTION AND		120
CSP=ec2		LAL CODECORPT	:	97
Bmar(SP2		ORD CTARLEY TY		118
accP11532		SSP CHOLGROLK	;	117
agor11002 p#70970v	1	REPORT AND A CONTRACT	:	124
SAP1		DAVED V - OP TO	:	105
SAP	2	COLOGY VALUE AND A CONTRACT AND A CO	:	111
am <sup>*</sup> 110969	2		:	80
ager10500	:	- SWILLW	:	114
ager0014	1	Brown a class bad c cc cc cc class contracts concerced a concerced back of the second stars be a before a concerced by the second stars be a before a concerced by the second stars be a before a concerced by the second stars by the second stars by the second stars before a concerced by the second stars before a concerced by the second stars by the second st	:	132

**Fig. 3** Partial alignment of candidate *Anopheles* CSPs with the previously identified CSPs shows the surplus sequence of agCP11435 that causes its 3D-structural abnormality. All acronyms of the gene names refer to the protein names listed in Table 1.

structure. It would be very interesting to discuss the number of insect OBPs and CSPs (29 OBPs to 9 CSPs for Anopheles), because the numbers may be closely related to the functions. It has been shown that the previously identified OBPs are spatially distributed in insect olfactory sensilla, while CSPs are distributed in different tissues, including non-olfactory tissues. Pheromone-binding proteins (PBPs) are typical insect OBPs, which are highly specific. However, CSPs are less specific, which indicate that one CSP may be able to bind more than one odorant. We postulated that more OBPs might be needed to bind different odorants specifically, although any further hypothesis on the physiological function of CSPs can only be made when reliable experimental evidence has been presented, such as identification of the specific ligands that show CSPs function as chemosensory proteins.

It must be noted that agCP11435 may be a good model for discussion on its structural relations to its functions. We hope agCP11435 is a CSP candidate, but its structural peculiarity may provide some space for exploring its binding behaviour. It is expected that a minor insertion into the structural core may lead to destruction of the fundamental functions of a protein.

No strong evidence has been provided for the physiological role of CSPs so far, though researchers believed they might be involved in chemical communication and perception. We checked the spatial expression pattern of the putative *Anopheles* CSP candidates (data unpublished) and found they were distributed not only in mosquito antennae (olfactory tissues), but also in non-olfactory tissues such as heads stripped off antennae and maxillary palps, legs and bodies. The preliminary results indicated that CSPs might have other functions than olfaction. Further studies will be focused on functional research, such as ligand identification and crystallization of CSP recombinants.

# Materials and Methods

# Defining conserved domains in insect CSPs

The amino acid sequences of previously identified CSPs were downloaded from GenBank. All the sequences were then aligned in ClustalX 8.1 (27) using Multiple Alignment Mode with the default gappenalty parameters. The multiple alignment was manually checked, and the absolutely conserved Cysteine residues in the alignment were defined as CSP motifs.

# Pattern search, selection of *Anopheles* CSP candidates in the whole genome sequences of *A. gambiae*

The Fasta files of gDNA, cDNA and the according translated peptide sequences of A. gambiae were downloaded from the Ensembl Mosquito Genome Browser (ftp://ftp.ensembl.org/pub/current\_mosqui to/). A pattern-searching program was created through standard PERL programming and named as CSPMOT by the authors. This program is run in Windows Commander, and can scan a local database for a pattern defined by the users. In this case, the CSP motif was used as a pattern to match every sequence stored in a local database that contains the Fasta files downloaded from the public databases. When a sequence matches the pattern, the program will display the sequence name, the position of the first residue where the pattern matches, and the pattern match alignment.

## BLAST

BLAST was performed at NCBI (http://ncbi.nlm. nih.gov/BLAST/; April 2003). The *Anopheles* CSP candidates identified by the program described above were used as queries to Blast GenBank. The E values smaller than 0.0001 were accepted. The E value is a parameter that describes the number of hits one can "expect" to see just by chance when searching a database of a particular size.

#### Annotations

GENSCAN (http://genes.mit.edu/GENSCAN.html; ref. 28, 29) and ORF Finder (http://www.ncbi.nlm. nih.gov/gorf/gorf.html) were used to predict the full-length genes, based on gDNA and cDNA sequences of Anopheles CSP candidates. The annotated CDSs were then used for calculation of hydrophobicity, pI and molecular weight of the CSP candidates, by using the comprehensive biosoftware Vector NTI (InforMax Inc., Bethesda, USA). The positions of signal peptides were determined online (http://www.cbs.dtu.dk/services/SignalP2.0/; ref. 30, 31), and finally the secondary structure, in particular the globularity, was predicted (http://maple.bioc.columbia.edu/predictprotein/; ref. 32-34).

# Homology modelling

Csp2 of Mamestra brassicae (PDB ID: 1K19; ref. 35) is an elucidated insect chemosensory protein, which was used as a model to predict the 3-D structure of Anopheles CSP candidates in Swiss-Pdb Viewer (http://us.expasy.org/spdbv/; ref. 36). Swiss-Pdb Viewer is a program that allows analysis of several proteins at the same time. It could be used to deduce 3-D structure of proteins and compare the active sites of different molecules.

# Acknowledgements

The authors thank Rothamsted International for its support of Dr. Zhengxi Li's post-doctoral fellowship in Biological Chemistry Division, Rothamsted Research, UK. Special thanks should be given to Division of Science and Technology, China Agricultural University for its financial support to the completion of the manuscript.

# References

- Holt, R.A., et al. 2002. The genome sequence of the malaria mosquito Anopheles gambiae. Science 298: 129-149.
- 2. Singh, R.N. and Nayak, S.V. 1985. Fine structure and primary sensory projections of sensilla on the maxillary palp of *Drosophila melanogaster* Meigen (Diptera:

Drosophilidae). Int. J. Insect Morphol. Embryol. 14: 291-306.

- Pelosi, P. and Maida, R. 1995. Odorant-binding proteins in insects. Comp. Biochem. Physiol. B. Biochem. Mol. Biol. 111: 503-514.
- Vogt, R.G., et al. 1999. Odorant-binding proteins diversity and distribution among the insect orders, as indicated by LAP, an OBP-related protein of the true bug Lygus lineolaris (Hemiptera, Heteroptera). Chem. Senses 24: 481-495.
- McKenna, M.P., et al. 1994. Putative Drosophila pheromone-binding proteins expressed in a subregion of the olfactory system. J. Biol. Chem. 269: 16340-16347.
- Pikielny, C.W., et al. 1994. Members of a family of Drosophila putative odorant-binding proteins are expressed in different subsets of olfactory hairs. Neuron 12: 35-49.
- Danty, E., et al. 1998. Separation, characterization and sexual heterogeneity of multiple putative odorantbinding proteins in the honeybee Apis mellifera L. (Hymenoptera: Apidea). Chem. Senses 23: 83-91.
- Briand, L., et al. 2002. Characterization of a chemosensory protein (ASP3c) from honeybee (Apis mellifera L.) as a brood pheromone carrier. Eur. J. Biochem. 269: 4586-4596.
- Maleszka, R. and Stange, G. 1997. Molecular cloning by a novel approach, of a cDNA encoding a putative olfactory protein in the labial palps of the moth *Cactoblastis cactorum. Gene* 202: 39-43.
- Robertson, H.M., et al. 1999. Diversity of odorantbinding proteins revealed by an expressed sequence tag project on male *Manduca sexta* moth antennae. *Insect Mol. Biol.* 8: 501-518.
- Picimbon, J.F., et al. 2000. Purification and molecular cloning of chemosensory proteins from Bombyx mori. Arch. Insect Biochem. Physiol. 44: 120-129.
- Nagnan-Le Meillour, P., et al. 2000. Chemosensory proteins from the proboscis of Mamestra brassicae. Chem. Senses 25: 541-553.
- Jacquin-Joly, E., et al. 2001. Functional and expression pattern analysis of chemosensory proteins expressed in antennae and pheromonal gland of Mamestra brassicae. Chem. Senses 26: 833-844.
- Picimbon, J.F., et al. 2001. Identity and expression pattern of chemosensory proteins in *Heliothis virescens* (Lepidoptera, Noctuidae). Insect Biochem. Mol. Biol. 31: 1173-1181.
- Pikielny, C.W., et al. 1994. Members of a family of Drosophila putative odorant-binding proteins are expressed in different subsets of olfactory hairs. Neuron 12: 35-49.
- Dyanov, H.M. and Dzitoeva, S.G. 1995. Method for attachment of microscopic preparations on glass for in situ hybridization, PRINS and in situ PCR studies.

BioTechniques 18: 822-824.

- Biessmann, H., et al. 2002. Isolation of cDNA clones encoding putative odourant binding proteins from the antennae of the malaria-transmitting mosquito, Anopheles gambiae. Insect Mol. Biol. 11:123-132.
- Angeli, S., et al. 1999. Purification, structural characterization, cloning and immunocytochemical localization of chemoreception proteins from Schistocerca gregaria. Eur. J. Biochem. 262: 745-754.
- Picimbon, J.F., et al. 2000. Chemosensory proteins of Locusta migratoria (Orthoptera, Acrididae). Insect Biochem. Mol. Biol. 30: 233-241.
- Marchese, S., et al. 2000. Soluble proteins from chemosensory organs of Eurycantha calcarata (Insecta, Phasmatodea). Insect Biochem. Mol. Biol. 30: 1091-1098.
- Normura, A., et al. 1992. Purification and localization of p10, a novel protein that increases in nymphal regenerating legs of *Periplaneta americana* (American cockroach). Int. J. Dev. Biol. 36: 391-398.
- Kitabayashi, A.N., et al. 1998. Molecular cloning of cDNA for p10, a novel protein that increases in nymphal regenerating legs of *Periplaneta americana* (American cockroach). *Insect Biochem. Mol. Biol.* 28: 785-790.
- Campanacci, V., et al. 2001. Chemosensory protein from the moth Mamestra brassicae. Expression and secondary structure from 1H and 15N NMR. Eur. J. Biochem. 268: 4731-4739.
- 24. Vogt, R.G., et al. 2002. A comparative study of odorant-binding protein genes: differential expression of the PBP1-GOBP2 gene cluster in Manduca sexta (Lepidoptera) and the organization of OBP genes in Drosophila melanogaster (Diptera). J. Exp. Biol. 205: 719-744.
- Galindo, K. and Smith, D.P. 2001. A large family of divergent *Drosophila* odorant-binding proteins expressed in gustatory and olfactory sensilla. *Genetics* 159: 1059-1072.
- Graham, L.A. and Davies, P.L. 2002. The odorantbinding proteins of *Drosophila melanogaster*: annotation and characterization of a divergent gene family. *Gene* 292: 43-55.
- Thompson, J.D., et al. 1997. The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25: 4876-4882.
- Burge, C. and Karlin, S. 1997. Prediction of complete gene structures in human genomic DNA. J. Mol. Biol. 268: 78-94.
- Burge, C.B. and Karlin, S. 1998. Finding the genes in genomic DNA. Curr. Opin. Struct. Biol. 8: 346-354.
- Nielsen, H., et al. 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Engineering 10: 1-6.

Geno., Prot. & Bioinfo. Vol. 1 No. 4 November 2003

- 31. Nielsen, H. and Krogh, A. 1998. Prediction of signal peptides and signal anchors by a hidden Markov model. In Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology (ISMB 6), pp.122-130. AAAI Press, California, USA.
- Rost, B. 1996. Predicting one-dimensional protein structure by profile-based neural networks. *Methods Enzymol.* 266: 525-539.
- Rost, B. and Sander, C. 1993. Prediction of protein secondary structure at better than 70% accuracy. J. Mol. Biol. 232: 584-599.
- 34. Rost, B. and Sander, C. 1994. Conservation and pre-

diction of solvent accessibility in protein families. *Proteins* 20: 216-226.

- Campanacci, V., et al. 2001. Recombinant chemosensory protein (CSP2) from the moth Mamestra brassicae: crystallization and preliminary crystallographic study. Acta Crystallogr. D. Biol. Crystallogr. 57: 137-139.
- Guex, N. and Peitsch, M.C. 1997. SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modelling. *Electrophoresis* 18: 2714-2723.