Cryptic Epitopes Induce High-Titer Humoral Immune Response in Patients with Cancer

Eliane Fischer, Sebastian Kobold, Sascha Kleber, Boris Kubuschok, Erik Braziulis, Alexander Knuth, Christoph Renner and Andreas Wadle

*J Immunol* 2010; 185:3095-3102; Prepublished online 26 July 2010;
doi: 10.4049/jimmunol.0902166
http://www.jimmunol.org/content/185/5/3095

**Supplementary Material**
http://www.jimmunol.org/content/suppl/2010/07/23/jimmunol.0902166.6.DC1

**References**
This article cites 51 articles, 23 of which you can access for free at:
http://www.jimmunol.org/content/185/5/3095.full#ref-list-1

**Subscription**
Information about subscribing to *The Journal of Immunology* is online at:
http://jimmunol.org/subscription

**Permissions**
Submit copyright permission requests at:
http://www.aai.org/About/Publications/JI/copyright.html

**Email Alerts**
Receive free email-alerts when new articles cite this article. Sign up at:
http://jimmunol.org/alerts
Cryptic Epitopes Induce High-Titer Humoral Immune Response in Patients with Cancer

Eliane Fischer,* † Sebastian Kobold,* Sascha Kleber,* Boris Kubuschok,‡ Erik Braziulis,§ Alexander Knuth,* Christoph Renner,* † and Andreas Wadle* †

In search of novel markers for diagnosis, prognosis, and therapy of cancer, screening of rcDNA expression libraries with patient’s sera has been established as a valuable tool for identification of cancer-specific Ags. Interestingly, besides the expected humoral responses to annotated proteins, patients with cancer were frequently found to have serum Abs that bind to peptides without homology to known proteins. So far, the nature of these unconventional epitopes and their possible significance in tumor immunity have never been thoroughly investigated. In our study, we specifically analyzed humoral immune response toward such peptides in patients with pancreatic or breast cancer using yeast-displayed cDNA expression libraries derived from tumor tissue. A detailed analysis of the identified peptides revealed that they originated from translation of sequences outside annotated open reading frames and may derive from the use of alternative start codons or from DNA indel mutations. In several cases, the corresponding mRNA templates have a known association with cancer. In a final analysis, we were able to detect one of these tumor Ags in cancer tissue arrays by a selected Fab-Ag. We conclude that cryptic epitopes may elicit specific humoral immune responses in patients with cancer and thus play a role in immunologic surveillance. Due to the high prevalence of immune responses against some of the peptides, they may also be valuable markers for cancer diagnosis, prognosis, or therapy monitoring. The Journal of Immunology, 2010, 185: 3095–3102.

Patients with cancer exhibit humoral immune responses against Ags expressed in tumor tissue, detectable as high-titer IgGs in blood serum (1–6). Hence, cancer-specific autoantibodies have been exploited to identify novel targets for therapy (7), markers for more sensitive diagnostics of various cancers (8–10), or simply better understanding of tumor biology. Most prominently, the serological analysis of rcDNA expression libraries method led to the identification of >2000 annotated antigenic sequences that are collected in the Cancer Immunome Database (11) to date.

Originally based on prokaryotic expression, the serological analysis of rcDNA expression libraries method was adapted to a yeast display-based system to facilitate folding and allow posttranslational modification of the expressed Ags (12–14), but also for efficient presentation of small peptides (15). We previously reported the use of rAg expressed on yeast surface technology for the discovery of novel Ags associated with breast cancer (16) and for the detection of autoantibodies to known cancer-testis Ags in patients with pancreatic cancer (17).

Over the years, we have identified a significant number of clones recognized by patient sera that were not encoded by known or postulated open reading frames (ORFs) and discarded them as potential artifacts. There is a growing body of evidence also reported by other authors indicating that specific humoral responses to cryptic epitopes can occur that were generated from cloning of putatively untranslated sequences into the expression system (8, 9). Despite their potential usefulness as biomarkers for various cancerous diseases, such cryptic epitopes are usually referred to as mimotopes (i.e., structures that mimic a structurally unrelated Ag), rather than real epitopes, and are not further investigated (8, 9). In consequence, little is known about their potential biological relevance in cancer immunity, although they could hold valuable information for a more holistic understanding of tumor-host interaction.

Still, there are virtually no data on humoral immune responses against actually proven cryptic epitopes. One exception is a study that describes the serological analysis of a cDNA library from microsatellite instability (MSI)-positive colon cancer cell lines (18). In this report, a patient with known MSI-positive colon cancer had a frameshift mutation in the coding region of CDX2, one of the identified Ags, and concomitantly showed an anti-CDX2 humoral response. In the patient’s serum, Abs against both neopeptides caused by the frameshift mutation and native epitopes of CDX2 were present. This is the first proof that tumor-specific peptides generated by MSI are involved in antitumor immune responses, possibly causing the better prognosis associated with MSI (19), and may, therefore, be useful for the development of new therapeutic approaches.

Although genetic inactivation of DNA mismatch repair leading to MSI is an elegant mechanistic explanation for the emergence of cancer-specific neopeptides, the latter may play a more general role in cancer immunity. Based on our previous serological analyses, we expect them to be present in a wide range of tumors. Therefore, we set out to systematically analyze antigenic cryptic epitopes from yeast-displayed cDNA libraries of breast and pancreatic cancer and provide an in-depth analysis of these unconventional epitopes.
Materials and Methods

Patients and samples

This study was approved by the local ethic committee (Ethikkommission der Ärztekammer des Saarlandes) and performed in accordance with the Declaration of Helsinki. After informed consent, serum samples were collected from 113 patients with inoperable or operable adenocarcinoma of the pancreas. In a proportion of patients (76%), tumor samples were obtained from pancreaticoduodenectomy (Whipple operation) or pancreatic-ectomy at the Saarland University Medical School (Homburg, Germany) between January 1997 and November 2000. To evaluate the specificity of the Ab responses for patients with cancer, tumor sera were compared with 63 sera from age- and sex-matched healthy donors and 18 sera from patients with acute or chronic pancreatitis. Furthermore, tumor and serum samples were obtained from 98 patients with operable breast cancer upon the primary definitive operation (mastectomy: 56 cases; lumpectomy: 59 cases; reductive tumor surgery: 5 cases) at the Department of Gynecology, University of Saarland Medical School (Homburg, Germany) between June 2001 and May 2002. In addition, 30 sera from women with autoimmune disease and 56 sera from age-matched healthy women were collected at the Department of Internal Medicine I, University of Saarland.

cDNA library construction

Total RNA was extracted from 15 adenocarcinoma or 15 breast cancer samples using the guanidium-thiocyanate method as previously described (16). Isolating the Fast Track kit 2.0 (Invitrogen, Leiden, The Netherlands). The cDNA synthesis was performed by the use of the Matchmaker Library Construction & Screening Kit (BD Biosciences, San Jose, CA) following the manufacturer’s instructions, and a modified PYD1 Vector (Invitrogen) was generated for cDNA library expression. The cDNA library was cloned by homologous recombination in the Smal linearized vector followed by a cotransformation into the competent Saccharomyces cerevisiae strain EBY100. Cultivation of yeast and induction of peptide expression was done as previously described (14, 17, 20).

Immunofluorescence staining for cell sorting

Patient sera were diluted 1:50 and two times preabsorbed on a yeast suspension (24 h at 4˚C) and then pelleted (1500 rpm, 5 min). Supernatants were used for immunofluorescence staining. The yeast cDNA library was divided in aliquots and stained with 20 different pools containing five single sera (each preabsorbed sera at 1:100 final dilution). Yeast cells were pelleted and washed three times in PBS/Tween 0.1%. Autoreactive bound IgGs were detected by the use of an anti-human IgG Ab, directly conjugated with PE (Jackson ImmunoResearch Laboratories, Suffolk, U.K.) as previously shown (14, 17, 20). Single cells were sorted by flow cytometry (DakoCytomation/Moffo, Hamburg, Germany). Parameters for the sorting procedure were established by double staining of a yeast clone expressing NY-ESO-1 using an anti-His Ab and polyclonal patient sera. The procedure was performed as previously described (17).

Plasmid DNA isolation from monoclonal yeast culture

The isolation of plasmid DNA was performed using the GFX Micro Plasmid Prep Kit (GE Healthcare, Munich, Germany). In addition to the manufacturer’s protocol, cell lysis was improved by addition of glass beads (Sigma-Aldrich, Munich, Germany) to the cell pellet (in a ratio of 1:2) and vigorous vortexing (30 min before isolation protocol). Extraction, amplification, and electrophoresis were performed as previously described (14, 17, 20). Inserts were sequenced using the Sequitherm EXCEL II DNA Sequencing Kit (Epigenetic Biotechnologies, Madison, WI) and a LI-COR DNA Sequencing Kit (LI-COR Biosciences, Munich, Germany). Sequences were assigned with the National Center for Biotechnology Information BLASTN Database (Bethesda, MD) (21).

Serological analysis by flow cytometry

Detection of immunological response at a single sera level was performed by the use of 30,000 yeast cells displaying the selected peptides on their cell surface. All preabsorbed sera were diluted 1:100, and staining was performed as described (14, 17, 20). FACS analysis was done using the WinMDI 2.8 software. Ratios of the mean fluorescence of Ag expressing yeast and mock-transformed yeast were calculated for each serum. A serum was considered responsive when this ratio exceeded the mean fluorescence of the healthy control group’s ratio by >5 times the SD (22). The significance of the association of positive immune response with disease was determined by Fisher’s exact test.

Serological analysis by ELISA

Serum reactivity on synthetic peptides was assessed by ELISA as previously described (23). The sera were used in a series of different dilutions (1:50–20,000, 100 µl/well), and data were collected by measuring the absorption at 450 nm with a microplate reader (Wallac Victor 2, PerkinElmer, Wellesley, MA).

Selection of Fab-Abs to the novel identified Ags

Selection, purification, and assessment of specific Fab-Abs were entirely performed by standard methods (24). Briefly, a phage library was used expressing a large nonimmune, human Fab repertoire, containing 1.45 × 10^10 different Ab fragments. Phages (1.8 × 10^5) were preabsorbed (1 h, room temperature) in 2% nonfat dry milk PBS with streptavidin-coated paramagnetic beads (Dyna, Oslo, Norway) and subsequently incubated (1 h) with decreasing concentrations of the biotinylated peptide (Peptides and Elephants, Potsdam, Germany: 300, 100, 20, and 5 nM). Streptavidin-coated beads were then added and mixed with constant rotation (15 min, room temperature). After eight wash cycles with 0.1% Tween-PBS and two additional washes with PBS, bound phages were eluted with 100 mM triethylamine and neutralized with Tris-HCl (pH 7.2). Phages were used to infect Escherichia coli strain TG1 (30 min, 37˚C) and bacteria were grown overnight (30˚C) on agar. The diversity of the selected Abs was determined by DNA sequencing. To confirm the reactivity of the selected Fab-Abs, 100 µl was at a concentration of 10 µg/ml for yeast clone staining. All staining steps were done at room temperature, and each sample was washed twice with PBS/Tween 0.1%. Bound Fab-Abs were detected by a biotinylated anti-human Fab-Ab (1:200; Jackson ImmunoResearch Laboratories) and streptavidin-PE (1:300; Jackson ImmunoResearch Laboratories).

ELISA with phage and purified Fab-Abs

The specificity of individual binders displayed on phage surface and soluble Fab-Abs was assessed by ELISA with coated biotinylated peptides as described (25). Recombinant soluble Fab-Abs were purified from E. coli periplasmic fraction as shown earlier (24). Bound Fab molecules were detected by the murine anti-Myc Ab 9E10 (Roche, Mannheim, Germany) and an HRP-conjugated anti-mouse IgG (DakoCytomation). Bound phage particles were detected by the Ab M13 recognizing the p8 protein (Amersham Pharmacia Biotech, Uppsala, Sweden) and an HRP-conjugated anti-mouse IgG (DakoCytomation). Detection was performed using tetramethylbenzidine reagent (Sigma-Aldrich).

Fluorescence microscopy

Fab B1 was modified by 3’-terminal biotin tag allowing for in vivo biotinylation at the C-terminal end (24, 26). Biotinylated Fab’s were tetram erized using Strept-A-FITC (Jackson ImmunoResearch Laboratories) (24, 26). Prior to immunostaining, tissue arrays with 90 breast cancer tissues and 10 normal breast tissue controls (Cybrdi, Rockville, MD) were deparaffinized in xylene (four times, 2 min each), followed by 1 min washes in EtOH 100% (two times), EtOH 95%, EtOH 80%, EtOH 70%, and finally water. After blocking in PBS containing 2% BSA (Sigma-Aldrich) for 30 min, samples were incubated with the tetramerized Abs (1 h, room temperature) and then washed three times in PBS for 5 min. For nuclear staining, sections were incubated for 5 min in PBS containing 1 µg/ml Hoechst 33342 (Sigma-Aldrich) and then washed twice for 5 min in PBS. Finally, the samples were mounted with Dako mounting solution (DakoCytomation) containing 25 mg/ml DABCO antioxidant (Sigma-Aldrich). Fluorescence microscopy was performed using a Nikon Eclipse TE2000-U inverted microscope equipped with a CoolSnap (Photometrics, Tucson, AZ) camera. Images were processed with Adobe Photoshop 7.0 (Adobe Systems, Munich, Germany).

Identification of yeast-displayed epitopes for M026-Ag specific Fab-Ab

cDNA-expressing yeast libraries were cultivated and peptide expression induced. A total of 25–50 µl yeast culture (OD 2-3) and 20 µl magnetic beads (Dynabeads M-280 and streptavidin 112.06D, Invitrogen) were washed and surface blocked by milk powder/PBS suspension (4%, 6 h, 4˚C). For magnetic cell sorting, periplasmic fraction as shown earlier (24, 26). Bound Fab molecules were detected by the murine anti-Myc Ab 9E10 (Roche, Mannheim, Germany) and an HRP-conjugated anti-mouse IgG (DakoCytomation). Bound phage particles were detected by the Ab M13 recognizing the p8 protein (Amersham Pharmacia Biotech, Uppsala, Sweden) and an HRP-conjugated anti-mouse IgG (DakoCytomation). Detection was performed using tetramethylbenzidine reagent (Sigma-Aldrich).

Secreton and lysate staining

For nuclear staining, sections were incubated for 5 min in PBS containing 1 µg/ml Hoechst 33342 (Sigma-Aldrich) and then washed twice for 5 min in PBS. Finally, the samples were mounted with Dako mounting solution (DakoCytomation) containing 25 mg/ml DABCO antioxidant (Sigma-Aldrich). Fluorescence microscopy was performed using a Nikon Eclipse TE2000-U inverted microscope equipped with a CoolSnap (Photometrics, Tucson, AZ) camera. Images were processed with Adobe Photoshop 7.0 (Adobe Systems, Munich, Germany).

Identification of yeast-displayed epitopes for M026-Ag specific Fab-Ab

cDNA-expressing yeast libraries were cultivated and peptide expression induced. A total of 25–50 µl yeast culture (OD 2-3) and 20 µl magnetic beads (Dynabeads M-280 and streptavidin 112.06D, Invitrogen) were washed and surface blocked by milk powder/PBS suspension (4%, 6 h, 4˚C). For magnetic cell sorting, periplasmic fraction as shown earlier (24, 26). Bound Fab molecules were detected by the murine anti-Myc Ab 9E10 (Roche, Mannheim, Germany) and an HRP-conjugated anti-mouse IgG (DakoCytomation). Bound phage particles were detected by the Ab M13 recognizing the p8 protein (Amersham Pharmacia Biotech, Uppsala, Sweden) and an HRP-conjugated anti-mouse IgG (DakoCytomation). Detection was performed using tetramethylbenzidine reagent (Sigma-Aldrich).
at 30˚C. This process was repeated two times. From monoclonal yeast cultures, pYD1-cDNA sequencing was performed as described earlier.

Results

Generation and sorting of pancreatic and breast cancer cDNA libraries

Two cDNA libraries from pancreatic adenocarcinoma and breast cancer tissue samples were cloned into the modified yeast display vector pYD1Rec. This vector allows expression of cDNA products as Aga2p fusion proteins on yeast surface. The transformed yeast libraries yielded \( \sim 10^6 \) primary transformants (670,000 CFU/ml).

Randomly picked clones were analyzed by insert sequencing (data not shown) and the size of cloned fragments determined by cDNA restriction enzyme digests varied from 100–1700 bp (Fig. 1A). Next, the libraries were enriched for yeast clones recognized by autoreactive Abs in the sera of patients with pancreatic and breast cancer. In the current study, we aimed to selectively search for yeast clones that react with the sera of patients with cancer and display an epitope derived from a cryptic sequence on yeast surface. Therefore, translated Aga2p fusion proteins without an in-frame C-terminal Penta-His-tag were selected. Following this strategy, we deselect clones with a continuous reading frame from the Aga2p protein to the His-tag. Detector settings for flow cytometry were deduced using a His-tagged NY-ESO-1–expressing yeast clone stained with a reactive patient’s serum (Fig. 1B).

Upon fluorescence staining with serum pools, 0.07% of the total library was selected by high-speed cell sorting (Fig. 1C, boxed portion of left panel). Despite the fact that only a minority of all clones expresses a C-terminal His-tag, an equal fraction of these

![FIGURE 1. Screening of rcDNA expression libraries. A, The diversity of cDNA inserts was demonstrated by double digest EcoRI/XhoI, and sizes of cDNA inserts were checked by agarose gel electrophoresis (top panel, pancreatic cancer library; bottom panel, breast cancer library). B, Determination of flow cytometry parameters for single-cell sorting of cDNA libraries. A NY-ESO-1–expressing yeast clone was stained by anti-His (green) and with a reactive patient serum and an anti-human Fc-specific PE-conjugate (red). Single- or double-stained yeasts are shown by forward light scatter/side scatter of light (B, top panel, first column) and fluorescence dot plot (B, top panel, second, third, and fourth columns). Cell surface staining was confirmed by fluorescence microscopy (B, middle and bottom panels). Original magnification \( \times600 \). C, Using the above-determined sorting criteria, the single-stained, penta-His–negative yeast population in the boxed portion of the left panel (0.07% of total library) was separated as described.](http://www.jimmunol.org/)

![FIGURE 2. Identification of cancer-specific Ags by serial serological analysis. Immune reactivity of serum samples from patients with cancer (A), healthy donors (B), and control groups (C) to 12 yeast clones expressing cryptic epitopes. Serum reactivity is indicated as standard deviations from the mean of the healthy donor group’s response as determined by FACS analysis (GMean ratio of fluorescence intensity on Ag expressing and mock-transfected yeast clones). The criterion for a positive immune response (mean + 5 SDs) is denoted by a horizontal line. Level of significance was determined by Fisher’s exact test and is indicated by asterisks. \( p < 0.05; \, **p < 0.01; \, ***p < 0.001. \)](http://www.jimmunol.org/)
were selectively recognized by patients with cancer to respective
We were able to map the regions for 8 of the 12 cryptic epitopes that
characterization of cryptic epitopes
Fisher’s exact test. However, patients also exhibited strong im-
significantly higher than in the healthy control group, as determined
from control groups with inflammatory diseases exhibited IgG re-
tumor patients were reactive to the individual peptide-expressing
(Fig. 2, Table I, Supplemental Data 1). Up to 16.2% of sera from
Healthy clones, whereas no healthy donors and only very few patients
Epitopes retrieved by cDNA expression on yeast surface
Peptide Sequence
M004 NM_203304 MEX3D GGGGRRRGGGGGGGGGGGGGGGACPRRT*
M008 NM_020895 GRAMD1A GGGGRRRGRGGPCAPQT5AAPRQAQPCPAPLCPALPCAR
M009 Chromosome 4 Genomic GGGGGRRRVRRT*
M017 NM_004719 [SFRS2IP] antisense GAGGGGGGGSF*
M026 NM_004739 MTA2 GAGGGGGNHHQQQVQGQRLRF*
M046 NM_00128833 ZBTB4 GAGGGGARAGAGAGAGAGAGAGAGGRPR*
M056 Chromosome 2 Genomic GGPGAAPAAATFDASRLPYDFSPCCRGKPRGARRAPAPPHCRRRPLR*
M061 Introns PTPRG GGGGGCGNQKGTCXAN*
M081 NM_001130 AES GGGGGGPOADRERGDAAARRRQQPEPEPTAGAQRRPRRD*
P483.1 NM_001025076 CUGBP2 GESPLPFSK*
P483.2 NM_002189 IL15RA HGSSAGDLDQQQR*
P166 NM_005940 MMP-11 HGGTGGTVSSLYFCCIVFSRGWH*
Stop codons are indicated by asterisks.
clones was stained with the patient’s sera pool (0.07% of the total
library; Fig. 1C, boxed portion of right panel). This generally
lower serum reactivity of the clones without C-terminal His-tag
is expected, because a large part of the peptides displayed by these
clones will be only artificially expressed on the yeast (e.g., by cloning artifacts), but will not be present in cells.
A total of 600 clones from a breast and a pancreatic cancer cDNA
library were sorted, sequenced, and translated into corresponding peptide sequences of the Aga2p fusion proteins. The putative
peptide length expressed on yeast surface ranged between 11 and
250 aa and was confirmed by Western blot analysis of the Aga2p
fusion proteins (data not shown). A high fraction of the peptides
were identified as cryptic epitopes, or more precisely as peptide
sequences without homology to known proteins. They mainly
originated from cDNAs that were cloned out of frame to the linker
protein or from 3’ and 5’ untranslated region (UTR) sequences
fused to the Aga2p protein. Besides that, some epitopes were
generated by cloning of short genomic DNA fragments.
Identification of cancer-related Ags
To identify cryptic epitopes that are selectively immunogenic in
patients with pancreatic or breast cancer, serial serological anal-
ysis of 99 patients with pancreatic cancer and 63 age- and sex-
matched healthy donors or 98 patients with breast cancer and 56
age-matched healthy women were performed. Patients with acute
and chronic pancreatitis and autoimmune diseases were evaluated
as a second control group in addition to healthy individuals.
We found 12 peptides with a higher percentage of serological responses
in the patient pool than in the control groups, in which a positive re-
sponse exceeded the mean of the healthy donor group by >5 SDs
(Fig. 2, Table I, Supplemental Data 1). Up to 16.2% of sera from
tumor patients were reactive to the individual peptide-expressing
yeast clones, whereas no healthy donors and only very few patients
from control groups with inflammatory diseases exhibited IgG re-
actions. For the clones M009, M026, M483.2, and P166, the number
of positive responders within the group of patients with cancer was
significantly higher than in the healthy control group, as determined
by Fisher’s exact test. However, patients also exhibited strong im-
mune responses to some of the other candidates, albeit less fre-
quently. The surprisingly common immune responses of patients with
cancer to cryptic epitopes were further characterized.
Characterization of cryptic epitopes
We were able to map the regions for 8 of the 12 cryptic epitopes that
were selectively recognized by patients with cancer to respective
ReSeq mRNAs (Fig. 3). The sequences of the remaining four
epitopes were not attributable to ReSeq mRNAs but identified as
genomic sequences.
Next, we elucidated whether there are mechanisms that could
explain the existence of the displayed peptides, or at least parts
thereof, in human cells. In principle, cryptic epitopes may derive
from all levels of gene and protein expression, including the use
of alternative reading frames, ribosomal frameshifting, or DNA mu-
tations (Table II). Four cryptic epitopes originated from the fusion of
a 5’ UTR sequence to the Aga2p protein. We therefore examined
whether there are any upstream ORFs that could give rise to these
products. Indeed, we found alternative start codons, CUG and ACG,
in frame with the expressed peptides in M008 and M026 (Supple-
mental Data 2). Notably, the ACG triplet that was found upstream
of the M026 sequence was in a strong Kozak context. The putative
use of this alternative start codon results in the translation of a 65-aa
peptide, and the C-terminal 14 aa are a −1 frameshifted product of
the annotated ORF for metastasis-associated protein 2 (MTA2).
Another mechanism at the translational level that could poten-
tially contribute to the expression of cryptic epitopes is ribo-
sonal frameshifting, which can occur on prone sites. We searched
for the corresponding characteristics of the cryptic epitope’s
mRNA for the clones that were translated in alternative reading
frames (P483.1, M004) by using the KnotInFrame algorithm to

Table I. Cryptic epitopes retrieved by cDNA expression on yeast surface

<table>
<thead>
<tr>
<th>cDNA Clone</th>
<th>GenBank Number</th>
<th>Identity</th>
<th>Peptide Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>M004</td>
<td>NM_203304</td>
<td>MEX3D</td>
<td>GGGGRRRGGGGGGGGGGGGGGGACPRRT*</td>
</tr>
<tr>
<td>M008</td>
<td>NM_020895</td>
<td>GRAMD1A</td>
<td>GGGGRRRGRGGPCAPQT5AAPRQAQPCPAPLCPALPCAR</td>
</tr>
<tr>
<td>M009</td>
<td></td>
<td>Genomic</td>
<td>GGGGGRRRVRRT*</td>
</tr>
<tr>
<td>M017</td>
<td>NM_004719</td>
<td>[SFRS2IP] antisense</td>
<td>GAGGGGGGGGGSF*</td>
</tr>
<tr>
<td>M026</td>
<td>NM_004739</td>
<td>MTA2</td>
<td>GAGGGGGNHHQQQVQGQRLRF*</td>
</tr>
<tr>
<td>M046</td>
<td>NM_00128833</td>
<td>ZBTB4</td>
<td>GAGGGGARAGAGAGAGAGAGAGAGGRPR*</td>
</tr>
<tr>
<td>M056</td>
<td>Chromosome 2</td>
<td>Genomic</td>
<td>GGPGAAPAAATFDASRLPYDFSPCCRGKPRGARRAPAPPHCRRRPLR*</td>
</tr>
<tr>
<td>M061</td>
<td>Introns</td>
<td>PTPRG</td>
<td>GGGGGCGNQKGTCXAN*</td>
</tr>
<tr>
<td>M081</td>
<td>NM_001130</td>
<td>AES</td>
<td>GGGGGGPOADRERGDAAARRRQQPEPEPTAGAQRRPRRD*</td>
</tr>
<tr>
<td>P483.1</td>
<td>NM_001025076</td>
<td>CUGBP2</td>
<td>GESPLPFSK*</td>
</tr>
<tr>
<td>P483.2</td>
<td>NM_002189</td>
<td>IL15RA</td>
<td>HGSSAGDLDQQQR*</td>
</tr>
<tr>
<td>P166</td>
<td>NM_005940</td>
<td>MMP-11</td>
<td>HGGTGGTVSSLYFCCIVFSRGWH*</td>
</tr>
</tbody>
</table>

FIGURE 3. cDNA sequences of cryptic epitopes mapped to mRNAs.
Schematic diagram of the location of the cryptic epitope (red boxes) in relation to the known ORF (white boxes) of each mRNA.
detect ~1 programmed frameshifts. However, we did not find evidence for ribosomal frameshifting to occur in these cases.

DNA mutations present a second source of cryptic epitopes by causing frameshifts by insertion or deletion of nucleotides or by affecting the genetic architecture (e.g., splice sites). We searched the Ensembl database for known single nucleotide polymorphisms on cDNA sequences. Interestingly, two sequences coding for cryptic epitopes were derived from mRNA regions with several identified indel mutations (P483.1, P166). The cDNA sequence of clone M081 also had deletion mutations. However, because all three sequences are not overlapping the ORF, the consequence of these mutations is unclear.

Finally, we found that some of the peptides contained repetitive, glycine-rich sequences. The immune reaction to these epitopes may be relatively unspecific. Thus, we classified them as mimotopes that may not be specifically recognized by patient’s sera, but instead mimic the structure of unrelated epitopes.

Next, we characterized the most immunogenic candidates, P166 and M026, in more detail (Fig. 4A, 4B). To this end, we used synthetic peptides with overlapping sequences for each Ag and tested binding of the highest responsive sera by ELISA. The C-terminal part of Ag P166 (VSLLYPCCEVPSRGW) was shown to contain the epitope recognized by the patient’s sera, whereas no reactivity was observed for the N-terminal part HGRTGGTGVSLYPC (data not shown). Likewise, the peptide covering the C-terminal part of Ag M026 (GHGGQHVPGGRLRLF), but not the N-terminal, glycine-rich sequence (GAGGGGGNGHGGQH), was recognized by the highest responsive breast cancer patient’s sera (data not shown). Down to a 1000-fold dilution of an immune responsive pancreatic cancer patient serum on yeast clone P166 (●), and a nonresponsive serum on yeast clone P166 (○), was recognized by the highest responsive breast cancer patient’s sera (data not shown). To this end, we incubated the original library with the Fab-Ab and selected yeast clones by performing two rounds of selection. With this procedure, we enriched several additional clones, and all of these clones shared a common sequence at the C terminus of the displayed peptide (Table III, Supplemental Data 1). We therefore conclude that the binding of B1 to M026 is sequence-specific, and the epitope most probably includes the sequence GRLRL.

In vivo biotinylated B1 was tetramerized by streptavidin A and used for staining of a tissue array containing mammary gland carcinoma and healthy tissues by immunofluorescence. A cytoplasmic staining pattern could be observed in 4 out of 90 carcinoma tissues, whereas all normal tissue samples were negative (Fig. 5B, 5C). These results indicate the presence of the M026 epitope in human breast cancer tissue.

**Discussion**

In our thorough analysis of cDNA expression libraries with patient’s sera, the strongest and most prevalent immune reactions were observed toward epitopes lacking homology to any known proteins. Such epitopes are the products of sequences outside annotated ORFs (e.g., 3’ and 5’ UTRs of the mRNA or alternative

**Detection of M026**

To detect cryptic epitopes in biological material, Ab Fab fragments specific for these products were selected by phage display technology. After three selection rounds, we isolated Fab-Ab B1, which specifically bound to the biotinylated peptide GHGGQHVPGGRLRLF (Fig. 5A). We tested whether the B1 Ab specifically binds to the GHGGQHVPGGRLRLF peptide sequence or if it is able to bind to other sequence-unrelated peptides displayed on the yeast surface. To this aim, we incubated the original library with the Fab-Ab and selected yeast clones by performing two rounds of selection. With this procedure, we enriched several additional clones, and all of these clones shared a common sequence at the C terminus of the displayed peptide (Table III, Supplemental Data 1). We therefore conclude that the binding of B1 to M026 is sequence-specific, and the epitope most probably includes the sequence GRLRL.

*We next compared the patient’s immune response to the cryptic epitopes M026, P483.2, and P166 with the serum reactivity to the in-frame translated products. To this aim, we mutated the corresponding cDNAs for in-frame expression in yeast. Interestingly, six patients with pancreatic cancer indeed had reactive sera for the in-frame expressed product of P166 (Fig. 4C), one patient to IL15RA (P483.2) (Fig. 4D), whereas no reactive serum was identified for correctly expressed MTA2 (M026) in patients with breast cancer (data not shown).*
reading frames as a direct result of the cDNA cloning procedure). This phenomenon is often explained with the occurrence of mimotopes (i.e., peptides that mimic an Ag [e.g., a tumor-specific glycosylation pattern] but are not homologous at the sequence level) (9, 10). In consequence, neither the origin nor the potential biological relevance of these peptides has been investigated further. Although we do not categorically exclude this hypothesis, we provide in this study evidence that at least some cryptic epitopes could actually be expressed in cancer tissue.

From the pancreatic cancer library, we isolated clone P166 displaying a peptide derived from the mRNA of matrix metalloproteinase 11 (MMP-11). Transcription of MMP-11 is significantly increased in pancreatic cancer tissue, where the protein is expressed on both tumor epithelial cells and the adjacent stroma (27, 28). A second example of a cancer-related Ag is a peptide derived from a frameshifted sequence of the MTA2. Overexpression of the latter leads to estrogen-independent growth of breast cancer cells (29). The cryptic epitope putatively arises by use of an upstream alternative start codon. We found a high incidence of humoral immune responses in patients with breast cancer to this Ag with 16.2% positive serum samples. We selected a Fab-Ab specific for the cryptic epitope that was able to stain breast cancer tissue but not normal tissue. In-frame cloning of the MTA2 protein, however, did not reveal autoimmunity of the patients with cancer to the correctly expressed protein.

Furthermore, we identified sera autoreactivity to a cryptic epitope mapped to the mRNA of a 5' UTR splice variant of the zinc finger and BTB domain containing protein 4 (ZBTB4). ZBTB4 binds methylated CpGs, thereby repressing gene transcription (30), and plays a crucial role in oncogenesis (31). In our cDNA clone, a CpG-rich sequence was alternatively spliced into the 5' UTR, and the immunogenic peptide M046 with the highly repetitive sequence 5'–GGAGGARAGAGAGAGGGQNRSGRGR–3' was derived thereof. Alternative 5' UTRs play an important role in posttranscriptional regulation, particularly in carcinogenesis (32, 33). Cryptic epitopes derived from 5' UTRs are a unique possibility to provide the immune system with a danger signal in response to disease-associated deregulation of gene expression. Notably, several other peptide sequences in our analysis are rich in amino acids Ala, Arg, Gly, and Pro and are derived from GC-rich nucleotide sequences, including 5' UTRs.

Each cell, healthy or diseased, produces not only functional proteins from the correct ORF, but also concomitantly gives rise to aberrant translational products through various mechanisms. For example, ≤30% of newly synthesized polypeptides have been shown to be defective ribosomal products resulting from premature termination or improper folding of the nascent protein (34, 35). A recent study of small proteins (<20 kDa) in cell lines demonstrated that the use of alternative translation start sites may significantly enlarge the human proteome (36). Numerous transcriptional and translational molecular mechanisms have been summarized to give rise to peptides expressed in alternative reading frames (37, 38). In short, there is accumulating evidence

Table III. Alignment of the C-terminal sequences of yeast-displayed peptides cross-reacting with B1-Ab and the M026 epitope

<table>
<thead>
<tr>
<th>Clone</th>
<th>C-Terminal Peptide Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>M026</td>
<td>HGGQHVPGGRRLFL*</td>
</tr>
<tr>
<td>B1-107</td>
<td>HAPPHV–GRGLIL*</td>
</tr>
<tr>
<td>B1-19</td>
<td>VVRPHV–GRGLIL*</td>
</tr>
<tr>
<td>B1-127</td>
<td>WALEAV–AAILLNK*</td>
</tr>
</tbody>
</table>

*After two selection rounds using Fab-Ab B1, three clones were enriched in the yeast library.

*Terminal sequences of the clones have been aligned using ClustalW. The putative binding motif is shown in boldface. Stop codons are indicated by asterisks.

**FIGURE 5.** Detection of a cryptic epitope in cancer tissue using an rAb-Fab fragment. Selection of an M026-specific Ab and detection of M026 expression in breast carcinoma tissue. A, Specificity of the selected Fab-Ab B1 was analyzed by ELISA. Nine different biotinylated peptides were coated on plates and reactivity measured by A450. Only peptide 1 (GHGGQHVPGGRRLFL) was recognized by Ab B1. Subsequently, selected B1-Fab and control Fab were biotinylated and tetramerized by streptavidin-FITC. Breast-cancer tissues (B) and healthy mammary gland control (original magnification ×400) (C) were stained by a nuclear dye (blue) and FITC-labeled tetramers (original magnification ×200). B and C, Overlay images of Hoechst and FITC are shown in the rightmost panels.
that a significant portion of the proteome consists of presumably short-lived, low abundant proteins that are not encoded by the annotated coding sequences.

What can be considered as molecular junk is actually a vast source of alternative antigenic structures to be monitored by the different branches of the immune system. Even if expressed only in minute amounts, such Ags can efficiently induce antiviral or anti-tumor T cell immunity by serving as MHC class I ligands. Indeed, many cryptic CD8 T cell epitopes have been characterized in viral infections (39–41), as well as on tumor cells (37, 42–47). MHC class II ligands have been less extensively studied in the context of cancer immunity. Nevertheless, CD4 T cell clones specific for off-frame translated MHC class II ligands have been reported (48, 49). In contrast, a general role of cryptic epitopes as B cell epitopes has not been described so far. Clearly, the presence of high-titer circulating tumor-specific IgGs reflects a more comprehensive immune response to an Ag, because it is dependent on concomitant T cell activation.

The question is how cryptic epitopes, despite their low abundance, small size, and probably short-lived nature, can actually elicit an efficient humoral immune response. The final identification and characterization of cryptic epitopes will always be hampered by their low abundance. In light of their significant impact on humoral immune response in patients with cancer, better understanding of the transcriptional, translational, and presentation mechanisms of these gene products is needed. Deeper insight into these concepts can finally facilitate the future design of vaccines and immunotherapeutic strategies against cancer (38, 50, 51).

Acknowledgments

We thank Birgit Bette for excellent technical assistance and Jochen Imig and Dan Hensel for helpful discussions and data analysis.

Disclosures

The authors have no financial conflicts of interest.
generated directly by transcription and translation of short subgenic regions. A hypothesis. *Immunogenetics* 29: 75–79.


Supplemental data 1

cDNA sequences of identified clones and the corresponding amino acid sequences of the displayed peptides. DNA/amino acid sequences of the Aga2p-protein are typeset in orange.

M004

```
L W P G G G G G R R G G G G R R G G G G S G
TTATGGCCGGGGAGAGGGGCGGCCGGGGACGGCGCGGCGGGGGGCGGCTCCGGAA
ACAPRRT*
GCTGCTGGACCCCAACGGAGATTCCTCCCGGCAGGGCAGGGCCCTGGGGTGCTG
```

M008

```
L W P G G G G G R R G R P G R C P A G T P S A A Q P
TTATGGCCGGGTGGGGGCGGCGGCCGCGGAAGGCCAGGCCGGTGCCCTGCGGGGACGCCCAGCGCAGCCCAGCCC
RARALPCPACPAPGARPP R R I H V R
CGGCGAGCCTGGCCTCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCCGGGCGGCGGCCCAACGCGGCACATCGTTGA
HHTPLWPEHAKQLPIAPETAAPAP
CACCACACCCACCTGGCGAGCAAGCAAGCAGCTCCCCATCGCTCCCGGAAACCGCTCGACTCTTGGCCCCCC
K P A P T *
AAGCCGGGCCCCACCTGAGCCAGAAAACCAGCCACCAGTGGAGGAGTGGCTAGTTAAGAAATCCGACTCCCG

M009

```
L W P G G G G G G R *
TTATGGCCGGGTGGAGGGCGGTGGAAGGGGAGGCAGGTGAGTGATATGGTTTGTCTGCAGCCCATTCAAgTCTCA
ACTCGWWTTaTRTCTCCCGGWtTTCCCATGTGCTGTAGGNGGGYCCCWGTGGGGGTaATTtTWTCGWGGGGCT
GCTTTTCTGCTGCTHTTCTCTGTGTTGTTTtCTGCTCTCTGTGGTTGTTTTTCNGGTTTCTGCTT
TTGCTCTTCTCCTCNTTTTTCGTTGCCCHCCCCgCNTGTC
```

M017

```
L W P G A G G G G G S P *
TTATGGCCGGGGGCTCGGGGGAAGGGGAGGGCTGGTAGTGGTGGATGTCTCCCTGAGATACACTGGGAGGAGTGGCTACCTGC
ATGGAGGGCACTCCTGTGAGATCGAAGGGGTAGCTGTAGGATGAAATTGAAATGGATTTGAGGTGCTATCTGCTCATC
AAAGAGCATGAACACCAACCATGCAATGGAAGATATTAGCCTAGCTGTTGTGCTGATTCTATTGTTGTTGCTGATT
ACAtttCAATTTTGCATTGAGATATTAGTGTGAAACCATCAACTGTGgtTGGTTTTGgtTCTTtTAG
GCTAGAAATCCTGTCAGAgtatttCCYCGCTtaTTTTCATCATCCTACTGACCTAGTGGAGCGTCAATTTTTTTGCTGATT
TCGTAAGTATAGTTTCTGACATCTCTCTCTGCGCAAAAGTCTCCTCCTGTTGTTTCTCCTGTCCTGTCCTGTC
TCTATCTGCTGGGTATACAAAGGAGAATCTCCTTGTCTCTTTTTCGTTATGGATTTCGNNCCACAGACCTGAA
```
TTATGGCCGGGGGCCGGAGGCGGCGGCGGGAACGGCCATGGCGGCCAACATGTACCGGGTGGGAGATTACGTCTA

TTTTGAGAACTCTTCCAGCAATCCTTACCTGGTTAGACGGATTGAGGAGCTCAACAAGACTGCAAATGGAAATGT

GGAGGAAAGTTGTGCTCTGTCTTTTCCGGGGCGAGGGAGTAAACCTAGCTGATGACTCTGACACACAGAGAAGGG

TTTCTTTCTCGGCAATTTGAATCATTACCGCAACCACACATACGGGGGAAATCGAGTGTGACCCCGCTGCAAGA

GACACTTCTCGTATCGAGCCGCGAGATTAGATGTGCTGAAATACCAAGCTGAGATCCCGTAAACAGCATCGAATTG

GGGAGAATCTgATAATCGGGAACCAgCAGAAGATGGAGATGAAGGTCTGGGACCCAGAACACCCTTccACA

TTTTCTTTCTCGGCAATTTGAATCATTACCCACACACATACGGGGGAAATCGAGTGTGACCCCGCTGCAAGA

GACAGATATACTCGAGGTGAATTAGATGTGCTGAAATACCAAGCTGAGATCCCGTAAACAGCATCGAATTG
TTTGCTATTTGTTAGAGTCTTTTACAMCATTTGTCTTCCAAAACTC
M061
L W P G G G G G A G N C K G T C K A N *
TTATGGCCGGGGGGAAGGAGGAGCGGGGAGGGGAGGACATGGGGTAAAGTGAAGCAGTCTATTACATGACTATCTGTT
AGGTGCTTTTTTTGCTCCTGTTTTTTGATGAGGCTCAGCGCGCACTCGG
AACATGAGATACTACAGGTCCTACGAGGCGGACTGTTGACTCTCTG
TTGTGCCTcATACTTGTGTTGAGGAAATgAATCGTCTGTTGGAATAGCATTTCACATGAGACTCTAGGAGTGGGAG
TCATTCAGAGAATGGGCGTCTGCTTCAAGTTACCAACACCTgC

M081
L W P G A G G G G P G A A D R R R R D K G A A A R
TTATGGCCGGGGGCAGGGAGGGCGGCGGCGGCCCGGGAGCTGCCGATCGGCGCCGGGACAAAGGCGCGGCCGCCGG
R P Q Q P E P G R T A G A Q R A P R R D *
CGCCCCCAGCAGCGGGGACAGGGGCGAGCGCGCAGCCCAGCCGGCGATTTGACATGATGTTTCC
ACAAAGCAGGACATCGGCGCTCTCGCACTTCACTCCCAAGGAACACTGACCTGACCATGAGCAGC
CAAAGCAGGAATTTCAGCTACTGCAAGcTCAGTACCACAGCCACAGTCGATATGTGACAGTTGACAGTGAAGA

P483.1
L W P G E S P F L P F S K K *
TTATGGCCGGGGGGAAGGAGGAGCGGGGAGGGGAGGACATGGGGTAAAGTGAAGCAGTCTATTACATGACTATCTG
CAGGGAGCTTTTTTTTTATCCTTTTCTCAGCAAGCAGAAGAAGACTGACGAGGGTGAAGGAGGACCACAGGAGATGAGACTTGGGAAATGCTCTCTTCATC
CACGGGAGGATCCTTTCTTTTCTTCTTCAGATGCAGAAGAATGTCATATTTTATAGGAGAATTGATGACAA
CACAGGAGGATCCTTTCTTTTCTTCTTCAGATGCAGAAGAATGTCATATTTTATAGGAGAATTGATGACAA
TGGCAGGATCTTCAAGGCGGCTACAGGTCATTTTCTTCTTCACCATAGGGGTTATAGGTGGCTCTGCTCAGAAT
ATCATTTTACTGTTTGTGTTTTATTATTCTTACTTGAATATTAAAAAAAATAAAATTTTAAAAAAACTGTGATT
TCATCTTTTATGAAATAGTCAATACAAAAATGTATTGTTTTTTTGTTGCTGTGAAGAATGTGATGAGATT

P483.2
L W P G H G S A G D L G D Q Q Q R *
TTATGGCCGGGGGCAATGAGGGCTCAGGCGGTGACTTTGCGGAGGAGACGCAAGAGATGAAGACTTGGGAAATGCTCTC
TCACCACCTATGAAACTCAGGGAGAACCAGCCAGCTAAGTCCGGAGTGGAGGACGCTCCCTCGGAGGAGGACGCTC
CGAAGTTGAGGAGGTGCAAGGAAAGCGGGCTTCAGGAGGACAGGCTACCCAGGCTGAGGATGGCTTGAGATG
CACGGACTGCGCGTGCCTCGCGGAGGAGCAGGGTTCTCAGCATTCCAGATGAGA

P166
L W P G H G R T G G T G V S L L Y P C C E V P S R
TTATGGCCGGGGGCAATGAGGGCTCAGGCGGTGACTTTGCGGAGGAGACGCAAGAGATGAAGACTTGGGAAATGCTCTC
CGAAGTTGAGGAGGTGCAAGGAAAGCGGGCTTCAGGAGGAGGACAGGCTACCCAGGCTGAGGATGGCTTGAGATG
CACGGACTGCGCGTGCCTCGCGGAGGAGCAGGGTTCTCAGCATTCCAGATGAGA

G W H *
Sequences of the three most abundant clones that were additionally identified after selection with Fab antibody B1

**B1-127 (37 clones)**

GGCTGGCACTGAAGCAAGGGTGCTGGGGCCCCATGGCCTTCAGCCCTGGCTGAGCAACTGGGCTGTAGGGCAGGGCCACTTCCTGAGGTCAGGTCTTGGTAGGTGCCTGCATCTGTCTGCCTTCTGGCTGACAATCCTGGAAATCTGTTC

**B1-107 (9 clones)**

**B1-10 (4 clones)**

GGCTGGC

CGAGACTGTAGCAGAGAAAAACGTATTTTATATTTCACAAAAA
CATTGAGTTTAACCCGCTGATCTGATAACACAGTGTAGATGTAACAAAATCGACTTTGTTCCACCTGTACTTT
TAGCTCGTACAAATACAATATAACTTTTCTATTTCTCCGTAACACATGTGTTCCTCCATGTAATATCCTTTTCTAT
TTTTGTTCCGTTACCAACTTTACACTTTATATAGCTATTTACTTCCTCTATACACTAAAAACTAAAGACAATT
TTAATTGGCTGCTGGCCATATTTTTCAATTTTGTATAAAATTTCCCTATAATTATCCTATAGTAGCTAAAAAAGAT
GAATGTGAATCGAATCCTAAGAGAATTGGCAAGTGCACAAACAATACTTAAATAAATACTACTCAGTAATAACCT
ATTTCCTAGCATTGTAGACGAAATTGCTATTTTG
Supplemental Data 2

Cryptic epitope sequences on ENSEMBL-Transcripts including genetic variation. Sequences of cryptic epitopes are shown in red, alternative start codons in blue.

Key

<table>
<thead>
<tr>
<th>Codons</th>
<th>Alternating codons</th>
<th>Alternating codons</th>
<th>Alternative start codons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variations</td>
<td>3' UTR</td>
<td>5' UTR</td>
<td>Frameshift coding</td>
</tr>
<tr>
<td>Other features</td>
<td>Non-synonymous coding</td>
<td>Regulatory region</td>
<td></td>
</tr>
<tr>
<td>Cryptic epitopes</td>
<td>UTR</td>
<td>cDNA-sequence</td>
<td>Peptide sequence</td>
</tr>
</tbody>
</table>

M004/ENST00000388824

1 ATGCCAGCTGCTGCTGGCCACCCGCCGCACGGGGCGGGCGGGCGGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG...
CATCTTCGTTG3GAGACCTTCTCTACCTAATCCTAGAGCCACACCGAGCCTTCACGCCATTAGCGCGGGCTCCTCTGCGAG

TGGAGAGACCACTGAGGCTCAAGCCGACCTGCTACACCTAACCTGAGCGCCTTCCAGGGCATTAGCCCGGGCCTCCTTGCGAG

TGAGAAGACACCCAATGGAGGCTACAAGCCCAAGCTCAATACACTCAAGCTCTACCGCCTGCTCCCCATGCGGGCAGCCAAGCGGCCCTAC

CAAGAAGCTACAGCGAGCGCCCGGAGGCTCTCCCTTCTCAACCTCACACGACACCCCGCCCGGGCCCTTCCAGGGCATTAGCCCGGGCCTCCT

TGGGCTCCACCTCCGGCACAGGCGTGGCCCTCCACAGCGCATACATTGCTCCACGAAGGGGCCCTTCTGCTCTGTTGCTATGAGGGGAGG

TAGCAAGTGGAGGGGGGAGGCTGACAGGGCGGAGGCTCCTGCTGCCGCTCCGCGGCGCCCTCTACCAGGAGGCTACTGCTGGTATC

CAAGAAAGCGAATACCACTACCTCTCCCCCCAGCGTGGCCCTCCACACCAGGAGCGCCCGGGCCCTTCTGCTACCTGAGGCACCGCCTCCG

GCCAGCCATGCTCCACACCAACCGAGGACAGGGGCTGGCCGGGGCCCCTCAGCTGCTGGCAGGCTCCACCTCCACTGTGTCAGATCATGC

GATTGGTGGAGGTGGGGGTCCCCCCACAGGGGCTGGCCGGGGCCCCTCAGCTGCTGGCAGGCTCCACCTCCACTGTGTCAGATCATGC

GCGAGCTCGTTGAGGGGGGAGGCTGACAGGGCGGAGGCTCCTGCTGCCGCTCCGCGGCGCCCTCTACCAGGAGGCTACTGCTGGTATC

CTACTCCAGAACCCCTAGAGCGAGCGGTCGCTCGATTGCGGAGGAGGCCCTTCTGCTCCCGCCGCTCCTCTGGGCGGGGCGGCGGCGG

GGGCTCCACCTCTGAGGAGCGCCCGGAGGCTCTCCCTTCTCAACCTCACACGACACCCCGCCCGGGCCCTTCTGCTACCTGAGGCACCGCCTCCG

TGCTGCTGAGCGGCCCGGGGGCACCCCAACCCCTGTCATTGCTATCGTGCAGGCTGTTGAGGAGGACACGTGGTCAGCTGGAGGGGAGG

GGCACTGGCGGCTCCACCGCCCTTCAACCCCGGGCCCTTCAACCCTGGGCTCCAGAAGGCGCAAGCTGCTGAGTCTGAGTCTGCAGGGCCCTCTTT

TGCTGCTGAGCGGCCCGGGGGCACCCCAACCCCTGTCATTGCTATCGTGCAGGCTGTTGAGGAGGACACGTGGTCAGCTGGAGGGGAGG

TGCGAGCACTGGCGGCTCCACCGCCCTTCAACCCCGGGCCCTTCAACCCTGGGCTCCAGAAGGCGCAAGCTGCTGAGTCTGAGTCTGCAGGGCCCTCTTT

CCAGGACCTTATCTATTGCGGAGGAGGCCACCCACATGTCGCAAGCGGGGGGCGGAGGACAGCTATGCTCTCCACCCAGACCCGGGGGACGTCAGAGGA

GCCCTTGATTGGGGCGCGCCCGGGAGAATCCGCGGTTCGAGGAGGAAATTCGGTGAGTGAGGCCGCGCCCGGACCGAT
```plaintext
TGTTGCTCTCTTGGCTAAAACTGGGCTCCCTGGCTCTGAGATTGAATATTTCCCATCCACCCACCTCCCCCCAAAAAAT
AGCTCACAAGGGGAGAGCCAGTATGGGGGAGCAAATTTGACAAATGGGAATTAGAGGAGTGCAGTTTTAAAAGGAAAAGTTGCTGTCATC
AAATGGCAGCCTTTTCCCCAGCTACTTCTTTGGGGCCAAGATGGCTGCCCTAGCAGCAATCACTGCCAAGGGCAAGATCATGGCTTTT
GGAGGGAGGTGAGTTTAGGGAGGGCCAGGACCATCCTCCTACCCCTCATACCCTCCCAGCATATACAAAAGGGGAGGTTTTAGACAGGCT
CCCTGAATGTTAACCACAGAGGAGTCACTCTTTACCTTCTCTCTCTTTGCACTTTTCTTGGTCTTGGCCACAGCCTGAGTGACGA
ATTTCCTACTGAATGTACCAAGTTCCAATTTTTAAGGGGGGGAAAGGTTTCAAATGGGGAAAAACACACAAAAAAAAAAATCACTAAAAA
TTCCCACAAATCTTGTTTCTGGCACTTTGAGAAACATTGCAAATAAACACTGCAAATGAAAATAAGAATACATATATATATATCTACACAAATTAT
ATATCTATCTATCTATACAGCGGAACCACAAGAGAGACTGAGGAAGGCCTGGAGGCAGGGGCAGAGGTGACGACAGTGCCCCTATATCCT
TAACCCATACTCCTGAGGCAAACAGGCATGGGAAAATGGGAAGGGTTGAGGATGGACCGGAGAATTGGAACTTCAGAATAGGTCAAAAT
TCCAAAACCATGGACATTTTTTTGGGAGAATTGAGATTGTAGACATTTTTTTTCTTAAATATGATCAAGGAAAATAGCTTCCAGAA
TGTGGTGGTTCTGGGCAACAAATGAGATTGTGGCGACGTGGAGATTAAAATATATGTATTTGAGCTGGGGAATTTGAATATTGTGAGTTT
CAGATGTTGGAAATTTGGGATTTTGCAACATTGTGCTTTTGGGCTGACTTTGGCCACAGCCTGAGTTCCCTGG
GAAGACGGGTGGTCAGAGTGAGAAGGCCACTGGTTCTGTGCCGCAGCACGCAAAATTTAGAATTCTACAGACTAGCTCTATACGTAGT
GAGGACCCAGATTTAGAAGAAACTGACCAATATTTATCTCCGCATTTGTGTGTGTCCAACTCTGTTGAGTTTT
TGAACTGTGCTCC
```

K
TTAAAAAACACACACACACCTGTGTATATAGCAGCTTTATCTCGTATCTCTACAGAGAGAGAGAAA

R
TGCTGTACAGGGGTGGCTAGCTGTGCCTTTCCAATAAGATGCTAGCTTC

7921 GTGTTTCTAGGGTATTTTTCTTTTTAAGTTTTGGATACAGGATTTTTTTCAATTAAATTATATGACTGCTTTATGACTIAAT

8011 ATTCAATATGCTGTTTGG

---------------------------------------------