RESEARCH PAPER

# Potential Aggregation-Prone Regions in Complementarity-Determining Regions of Antibodies and Their Contribution Towards Antigen Recognition: A Computational Analysis

Xiaoling Wang • Satish K. Singh • Sandeep Kumar

Received: 27 January 2010 / Accepted: 30 March 2010 / Published online: 27 April 2010 © Springer Science+Business Media, LLC 2010

# ABSTRACT

**Purpose** To analyze contribution of short aggregation-prone regions (APRs), which may self-associate *via* cross- $\beta$  motif and were earlier identified in therapeutic mAbs, towards antigen recognition *via* structural analyses of antibody-antigen complexes. **Methods** A dataset of 29 publically available high-resolution crystal structures of Fab-antigen complexes was collected. Contribution of APRs towards the surface areas of the Fabs buried by the cognate antigens was computed. Propensities of amino acids to occur in APRs and to be involved in antigen binding were compared. Coincidence between APRs and individual CDR loops was examined.

**Results** All Fabs in the dataset contain at least one APR in CDR loops and adjacent framework  $\beta$ -strands. The average contribution of APRs towards buried surface area of Fabs is  $16.0 \pm 10.7\%$ . Aggregation and antigen recognition may be coupled *via* aromatic residues (Tyr, Trp), which occur with high propensities in both APRs and antigen binding sites. APRs are infrequent in the heavy chain CDR 3 (H3) loops (7%), but are frequent in H2 loops (45%).

**Conclusions** Co-incidence of APRs with antigen recognition sites can potentially lead to the loss of function upon aggregation. Rational structure-based design or selection strategies are suggested for biotherapeutics with improved druggability while maintaining potency.

**Electronic Supplementary Material** The online version of this article (doi:10.1007/s11095-010-0143-5) contains supplementary material, which is available to authorized users.

X. Wang • S. K. Singh • S. Kumar (🖂) Pharmaceutical Research and Development BioTherapeutics Pharmaceutical Sciences Pfizer Inc. 700 Chesterfield Parkway West Chesterfield, Missouri 63017, USA e-mail: Sandeep.Kumar@Pfizer.com **KEY WORDS** biotherapeutics · drug development · fab-antigen interface · monoclonal antibody · structure

# **ABBREVIATIONS**

APRaggregation-prone regionCDRcomplementarity-determining regionFabfragment antigen bindingFcfragment crystallizableFRframework regionmAbmonoclonal antibodyPDBProtein Data Bank

# INTRODUCTION

Biotherapeutics, including monoclonal antibodies (mAbs) and their fragments, are an important segment of the pharmaceutical industry (1,2). Due to their ability to bind the targets with high specificity and affinity as well as near absence of non-mechanism toxicity due to off-target binding (3), mAbbased drugs offer attractive advantages over the small molecule therapeutics. On the other hand, these biotherapeutic drugs possess complex and conformationally heterogeneous molecular structures that are vulnerable to the changes in their environments and themselves. A complex series of processing steps between production to administration results in several physico-chemical stresses on the molecules (4). These stresses include high concentrations, variable temperatures, pH extremes, varying ionic strengths, shear stresses, and air-water as well as a variety of solid-liquid interfaces. As a result, potency and purity of the final drug product is impacted via multiple degradation pathways (5). Hence, there is considerable interest in the biopharmaceutical industry towards gaining fundamental understanding of the molecular properties that determine chemical, thermodynamic and long-term stability of biotherapeutic molecules.

Aggregation is the most common degradation pathway for biotherapeutics. Besides their potential to impact drug potency, aggregates are also considered a risk factor for immunogenicity (6). In particular, cross  $\beta$ -aggregates in biotherapeutics have the potential to be immunogenic (7). Hence, fundamental understanding as to why some molecules are more prone to aggregation than others would go a long way towards reducing or even eliminating this risk factor.

Aggregation is also a topic of intense research for proteins in general. Accumulating experimental evidence shows that specific regions of protein sequences, especially the ones with amyloidogenic properties, tend to drive aggregation (8-12). Early studies on small proteins and peptides have led to characterization of the physico-chemical properties of amyloid or amyloid-like aggregates associated with neurodegenerative diseases (13). The molecular trigger for these aggregates is the generation of the cross- $\beta$  motif whose molecular structure was recently elucidated by Eisenberg lab (14,15). Formation of cross- $\beta$  motif and amyloid-like aggregates in proteins is quite common. From a survey of the literature, we have found that experimental evidence is available for more than seventy different proteins showing aggregation via formation of the cross- $\beta$  steric zipper motif (16), and even proteins in bacterial inclusion bodies can aggregate via this route (17). Short sequence regions that potentially drive aggregation have been detected in these proteins. These are called aggregation prone regions (APRs) (18). Typically, these APRs have unique features with respect to charge, hydrophobicity, aromaticity and secondary structural preference. A number of computational approaches have been developed to predict potential APRs in proteins (16). Most of these prediction methods use only the protein sequences as input to identify short APRs of 5-9 residues capable of forming amyloid-like fibrils (19). Other methods based on pattern recognition, three-dimensional profiles and molecular simulations are emerging (20-26).

The question of whether there are also APRs in mAbs led to our hypothesis that non-covalent aggregation in biotherapeutics has many parallels with that seen in proteins in general. Thioflavin T and Congo Red are the marker dyes commonly used for detecting amyloid-like aggregation because they bind the cross- $\beta$  steric zipper motif (14). Aggregates formed by biopharmaceuticals, including therapeutic mAbs towards the end of their expiration dates, were reported to bind Thioflavin T and Congo Red (7,27).

Recently, we used TANGO (28) and PAGE (29) to identify the potential APRs in commercially available therapeutic mAbs (30). All therapeutic mAbs in our study contain several sequence regions that are strongly predicted to be aggregation prone (30). An interesting finding of our study is that some of these APRs are located in variable

domains, primarily in complementarity-determining regions (CDRs) and adjoining framework  $\beta$ -strands (30). These parts of the antibody molecule also contribute significantly towards antigen binding. Hence, the above study indicated that there may be an undesirable link between aggregation tendency and molecular function in the therapeutic mAbs.

Here, we present a statistical survey based on high resolution crystal structures of Fab-antigen complexes available in the Protein Data Bank (31). These also include Fabantigen complexes for commercially available therapeutic mAbs. We identify potential APRs in Fabs in our dataset using sequence-based prediction tools TANGO (28) and PAGE (29). The solvent-accessible surface area buried upon Fab-antigen complex formation is used as a parameter to gauge Fab-antigen interaction and the contribution of the potential APRs towards antigen recognition. In literature, protein-protein interactions have been commonly measured using approaches that rely on energetics (32,33) or contact residues (34) and the buried surface area (35) at the interfaces. The two approaches are closely related. In this study, we preferred to use buried surface area as a metric because it can be easily calculated. Moreover, it has been widely used in the study of protein-protein interactions (35-37).

All Fabs in our dataset contain at least one APR. In all but one Fab-antigen complex, the residues in these CDRlocalized APRs also contribute towards antigen binding. Hence, we find that potential APRs may contribute significantly towards antigen binding. The aromatic residues, Tyr and Trp, occur with high propensities in both APRs and antigen binding sites. Incidence of APRs is not uniform in all CDR loops. In particular, APRs are less frequent in the heavy chain CDR 3 (H3) loops. On the other hand, APRs are found with the greatest frequencies in the heavy chain CDR 2 (H2) loops. Rational structure-based design strategies for therapeutic antibody candidates with both high potency and improved druggability are proposed.

## MATERIALS AND METHODS

#### **Dataset of Fab-Antigen Complexes**

The term *antigen* is used throughout to refer to all binding partners of Fab in the complexes as defined by Janeway *et al.* (2004) (38). The complexes included in this study were collected from the Protein Data Bank (31) based on the following criteria. (a) The resolution is 2.5 Å or better. (b) Antigen is protein with at least 50 residues; complexes with small molecule haptens, polysaccharides, peptides and RNA as antigens are not included. (c) The complex has full Fab structure; complexes with Fv or single chain of antibody are not included. (d) The antibody sequences contained in complexes are non-redundant. This last criterion was hard to meet because the antibody sequences show high homologies when both variable and constant regions are included in the alignments. We retained one complex with best resolution from a set of complexes where both light chains and heavy chains of the Fabs showed greater than 90% sequence identity. The choice of 90% sequence identity is arbitrary. The average sequence identities for variable regions of Fab light and heavy chains in the final dataset are 60% and 55%, respectively. This indicates that we have a fair coverage of antibody sequence variation in our dataset. If more than one complex were present in the crystallographic asymmetric unit, only one copy is retained. (e) Catalytic antibodies are not included.

Our final dataset consists of 24 Fab-antigen complexes (resolution range: 1.2–2.5Å). We supplement this dataset with all the five available commercial Fab-antigen complexes. Among the five commercial Fab-antigen complexes, one complex (PDB ID: 1CE1) has an eight-residue-long peptide as antigen. Antigens in the other four complexes are proteins. These complexes satisfy the selection criterion on antigen type and size but not on resolution. The resolution for these structures varies from 1.9 Å to 2.61 Å. Overall, there are 29 complexes in our study with resolution range of 1.2–2.61 Å. This dataset compares favorably with the crystal structural datasets used in the previous studies of antibody-antigen complexes (32,35,36).

# **Definition of CDRs**

Andrew Martin's definition of complementarity-determining regions (CDRs) in antibody is followed (36). We chose this definition because it is based on antibody-antigen contact analysis, which shares similarity with our identification of recognition sites. Kabat numbering scheme is adopted (39).

#### Identification of Potential Aggregation-Prone Regions

The potential aggregation prone regions (APRs) are identified using a similar approach as described in our previous work (30). Two sequence-based prediction tools, *viz.* TANGO (28) and PAGE (29), are used to identify potential APRs. The aggregation propensity ( $\ln \pi$ ) from PAGE is converted to Z score to identify the regions with statistically high aggregation propensity. The Z score of residue *i* is calculated as follows:

$$Z_{i} = \frac{\ln(\pi_{i}) - \overline{\ln(\pi)}}{std(\ln(\pi))}$$
(1)

where  $\overline{\ln(\pi)}$  is the average aggregation propensity of the sequence, and *std(ln\pi)* is the standard deviation about average aggregation propensity.

We identify a region of sequentially consecutive residues as *aggregation prone* (APR) if any of the following criteria is satisfied:

- (i) The TANGO scores of five or more consecutive residues are  $\geq 10\%$ . Such APRs are assigned type 1 and are considered strong predictions.
- (ii) PAGE Z score is at least 1.96. Such APRs are assigned type 2 and are considered strong predictions.
- (iii) TANGO scores are  $\geq 5\%$  and PAGE Z score is  $\geq 1$ . Such APRs are assigned type 3 and are considered weak predictions.

TANGO was reported to yield a success rate of 92% for peptides with TANGO score of 5% or greater (28). We use a more stringent cutoff of 10% to ensure APRs of type 1 are strongly predicted. APRs of type 2 are also strong predictions, as Z=1.96 corresponds to 95% statistical confidence level. APRs of type 3 are weak predictions due to lower cut-off values.

As an example, Fig. 1 shows the TANGO and PAGE profiles for the light chain of the Fab in VEGF-blocking Fab—Neuropilin-1 complex (PDB ID: 2QQN). The types of APRs identified with the above criteria are labeled. This chain has all three types of APRs. APRs of type 3 are not frequent in our dataset. The APR at residue 131–136 is of both type 1 and type 2.

We deliberately used two computational programs with substantially different algorithm philosophies. This ensures that identification of APRs in our dataset is not influenced by the peculiarities of the training sets and algorithms used in the development of computer programs. We note that APRs of type 1 and type 2 often overlap in our Fab sequences. The overlapping APRs are merged into single APRs in our data analysis.

## Optimization of Antibody-Antigen Complex Structures

All structures of Fab-antigen complexes have been subjected to optimization. The optimization is performed using the modeling package Molecular Operating Environment (MOE) from Chemical Computing Group (http://www. chemcomp.com/). The optimization includes two major steps, *viz.*, building the missing heavy atoms or residues (if any) and energy minimization.

#### Step One: Building the Missing Heavy Atoms and Residues

Each complex structure is first scanned for missing heavy atoms or residues. Complexes without missing atoms or residues are optimized *via* energy minimization only, and this step is skipped.

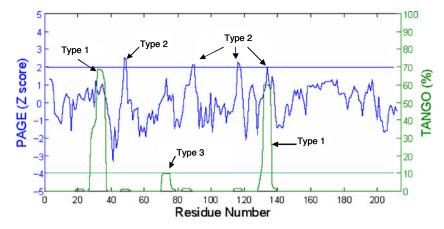


Fig. I The TANGO and PAGE profiles for the light chain of the Fab in VEGF blocking Fab—Neuropilin-I complex (PDB ID: 2QQN). X-axis shows residue number. Left Y-axis and blue curves are for PAGE Z score. Right Y-axis and green curve are for TANGO aggregation percentage. The blue horizontal line indicates PAGE Z score = 1.96. The green horizontal line indicates TANGO aggregation percentage = 10%. The peaks are labeled according to type of APRs (see Material and Methods).

If the missing atoms in the complex are side-chain-heavy atoms, they are built using the rotamer library in MOE. The qualities of the side-chain models built in MOE are ranked based on energies, RMSD, and side-chain torsion angles. The side-chain model with first rank is chosen.

For those complexes with missing backbone atoms or residues, the full structures are built *via* homology modeling. The chain containing missing residues is the target to be built. The original structure of the target is used as primary structure template in homology modeling. An additional structural template is chosen such that it shares high overall sequence identity with the target sequence and contains coordinates for the missing residues. Thus, the additional template is applied only for the missing residues. In each homology modeling, 100 intermediate models are generated and ranked by Generalized Born/Volume Integral (GB/VI) scoring (40). The model with the best GB/VI score is chosen and further optimized *via* energy minimizations.

#### Step Two: Energy Minimizations of the Complexes

All-atom AMBER99 force field is used in energy minimization of each complex (41). A cut-off of 12 Å with switching started at 10 Å is applied to van der Waals as well as Coulombic interactions. Generalized Born-implicit solvation is employed. Interior and exterior dielectric values are 4 and 80, respectively.

Energy minimization is performed in two steps: (a) the hydrogen atom positions are optimized while heavy atoms kept harmonically tethered, and (b) all atoms are then energy minimized until gradient falls below 0.001. The optimized structures show less than 2 Å C $\alpha$  RMSD with respect to the original structures.

#### Fab-Antigen Recognition Sites and Interface Areas

We use buried surface area to identify Fab-antigen recognition sites and gauge Fab-antigen interface. The buried surface area is referred to the surface area on both Fab and antigen that is accessible to solvent when Fab and antigen are separated but becomes inaccessible to solvent due to Fabantigen complex formation. The solvent-accessible surface area (ASA) is calculated using the algorithm of Lee & Richards as implemented in the program Accelrys Discovery Studio (42). The probe of water solvent is 1.4 Å in radius.

The total buried surface area of a complex can be simply obtained as the sum of the ASA values of its isolated components minus that of the complex. Here, we calculate buried surface area in an equivalent hierarchical way in order to obtain the following quantities: (a) residues in Fab-antigen recognition sites, (b) total buried surface areas of the complex and contributions from its Fab and antigen components, (c) contribution of APRs in Fab to buried surface area, and (d) polar fraction of buried surface area.

Our hierarchical calculations start at atom level. First, the buried surface area of an atom is obtained as the difference between the ASA of that atom in isolated component and in the complex. Second, the buried surface area of a residue is calculated as the sum of the buried surface areas of all atoms in that residue. The binding sites are identified at residue level. A residue is considered as a Fab-antigen recognition/binding site if its buried surface area is  $\geq 6$  Å<sup>2</sup>. The cut-off of 6 Å<sup>2</sup> is arbitrary. We have also tried lower cut-off values of 3–5 Å<sup>2</sup>. The differences are minor. Third, the buried surface areas of Fab and antigen are the sums of those of binding residues in Fab and antigen, respectively. In a similar way, buried surface area of APRs in Fab is the sum of those of binding residues which also belong to APRs. Finally, total buried surface area of the complex is obtained as the sum of buried surface areas of Fab and antigen. Polar and nonpolar buried surface areas are summed from polar and non-polar atoms, respectively. In our calculation, all carbon atoms are treated as non-polar. All oxygen and nitrogen atoms are considered polar. The sulfur atoms are considered non-polar if they are disulfide bonded. Otherwise, they are treated as polar.

We have also tried to identify the interface contact residues using the distance methods as employed by Tsai *et al.* (1996) (34). A residues pair is considered to be in contact between Fab and antigen if at least one heavy atom pair in the two residues is within a cut-off value (5Å). We obtained almost identical sets of binding site residues.

#### Hydrogen Bonds and Ion Pair Contacts

The program MOE is used to search hydrogen bond and ion pair contacts between Fab and antigen. The criteria for hydrogen bond proposed in Reference 43 is followed (43). The ion pair contact is inferred if a pair of nitrogen and oxygen atoms belonging to oppositely charged residues is within cut-off of 4.5 Å.

#### **Propensity Value Calculations**

Propensity  $P_{bi}$  of amino acid *i* to be a binding residue in Fab is calculated using the equation (44)

$$P_{bi} = \frac{n_{bi}/n_i}{\mathcal{N}_b/\mathcal{N}} \tag{2}$$

$$\mathcal{N}_b = \sum_{i=1}^{20} n_{bi} \quad \mathcal{N} = \sum_{i=1}^{20} n_i$$
 (3)

where  $n_{bi}$  = number of amino acid *i* at binding sites of Fabs,  $n_i$  = total number of amino acids *i* in the heavy and light chains of the 29 Fabs in our dataset,  $N_b$  = total number of binding residues in Fabs, and N = total number of amino acids in heavy and light chains of the 29 Fabs.

Propensity  $P_{ai}$  of amino acid *i* to occur at APRs in Fab is calculated in similar way, using the equation

$$P_{ai} = \frac{n_{ai}/n_i}{\mathcal{N}_a/\mathcal{N}} \tag{4}$$

$$\mathcal{N}_a = \sum_{i=1}^{20} n_{ai} \tag{5}$$

where  $n_{ai}$  = number of amino acid *i* in APRs of Fab,  $N_a$  = total number of APR residues, and  $n_i$  and N have same meaning as in Eqs. 2 and 3.

#### RESULTS

Our dataset contains 29 Fab-antigen complexes taken from the Protein Data Bank (Table I). Twenty-four of these structures (Number 1–24 in Table I) are for non-commercial Fab-antigen complexes selected based on criteria outlined in Materials and Methods. The bottom five complexes in Table I are the commercially available Fabantigen complexes. In all but one (PDB ID: 1CE1) case, the antigens in our dataset are proteins with at least 50 amino acid residues. Inclusion/exclusion of this complex does not impact our results significantly.

#### Location of Potential Aggregation-Prone Regions

The multiple sequence alignments of light chains and heavy chains of Fabs are presented in Fig. 2a and b, respectively. The potential APRs obtained from TANGO (28) and PAGE (29) analysis are highlighted in Fig. 2 to facilitate direct comparison among the complexes. The APRs in variable domains (V<sub>L</sub> and V<sub>H</sub>) are primarily located in CDR loops and adjoining  $\beta$  strands. The APRs in constant domain (C<sub>L</sub> and C<sub>H</sub>) are well conserved in terms of location and composition. Each Fab in the complex contains 3-9 APRs with at least one of them being in CDR loops. These observations are consistent with our previous finding (30), even though different criteria are used here (see Materials and Methods section for details). In this study, we classify the APRs into types 1, 2, and 3. We consider APRs of types 1 and 2 to be strong predictions, while the type 3 APRs are weak predictions. Most of the APRs found in the CDRs and adjoining regions are of types 1 and 2. The APR "FTLTI" located in light chain framework region 2 for a few Fabs is of type 3. However, this APR is not involved in antigen binding.

# Contribution of CDR Localized APRs Towards Antigen Recognition

Table II (column 3 and 4) lists the buried surface area of Fab and antigen. Total buried surface area of a complex is the sum of buried surface areas of its two components, namely, Fab and antigen. Total buried surface areas (column3 + column4) of the 29 complexes range from 900 to 3,000 Å<sup>2</sup>, reflecting the general nature of proteinprotein interfaces (35,37,45,46). Contributions from Fab and antigen to total buried surface area are close to halfand-half. The average total buried surface area for the complexes in our dataset is 1,915±504 Å<sup>2</sup>, if we exclude therapeutic Fab-peptide complex (PDB ID: 1CE1). The buried surface area in this excluded complex is particularly small (926 Å<sup>2</sup>) due to the small size of peptide antigen.

Table I List of Fab-Antigen Complexes

Number	PDB ID	Description	Resolution (Å)	R-Free
I	I FE8	Von Willebrand factor A3 domain/Fab fragment of IGG RU5 that inhibits collagen binding	2.03	0.264
2	I FNS	Von Willebrand factor A1 domain I546V mutant/the function blocking Fab NMC4	2	0.207
3	1H0D	Human angiogenin/Fab of mAb 26-2F	2	0.272
4	IIQD	Human factor VIII C2 domain/human monoclonal BO2C11 Fab.	2	0.253
5	IJPS	Tissue factor/humanized Fab D3h44	1.85	0.224
6	IKB5	Murine T-cell variable domain/Fab	2.5	0.221
7	ILK3	Engineered human interleukin-10 monomer/9D7 Fab fragment	1.91	0.24
8	IMLC	FAB D44.1/lysozyme	2.5	0.282
9	IOSP	Outer surface protein A of borrelia burgdorferi/Fab of a murine mAb	1.95	0.295
10	I UJ3	A humanized Fab fragment of anti-tissue-factor antibody/tissue factor	2.1	0.227
11	I WEJ	IgG1 Fab fragment (of E8 antibody)/horse cytochrome C	1.8	0.256
12	IYQV	Fab HyHEL5/lysozyme	1.7	0.234
13	IZTX	West Nile virus envelope protein DIII/neutralizing E16 antibody Fab	2.5	0.282
14	2B2X	VLA1 RdeltaH I-domain/a quadruple mutant of the AQC2 Fab	2.2	0.272
15	2CMR	HIV-1 neutralizing antibody D5 Fab/the GP41 innter-core mimetic 5-helix	2	0.258
16	2DD8	SARS-CoV spike receptor-binding domain/neutralizing antibody	2.3	0.261
17	2FD6	Human urokinase plasminogen activator/urokinase receptor and an anti-upar antibody	1.9	0.276
18	2NXY	HIV-1 gp120 envelope glycoprotein(S334A)/CD4 and antibody 17b	2	0.231
19	2Q8B	Malaria antigen AMA1/growth-inhibitory antibody	2.3	0.256
20	2QQN	Neuropilin-1 b1 Domain/VEGF-blocking Fab	2.2	0.207
21	2R0L	Short form HGFA/Inhibitory Fab75	2.2	0.248
22	2VDR	Integrin alphallBbetaA3 headpiece/a chimeric fibrinogen gamma chain peptide	2.4	0.193
23	3D85	Crystal structure of IL-23/neutralizing Fab	1.9	0.214
24	3D9A	HyHeIIO Fab/hen egg lysozyme	1.2	0.205
25	I BJ I	Vascular endothelial growth factor/neutralizing antibody	2.4	0.266
26	ICEI	Therapeutic antibody Fab/a synthetic peptide antigen	1.9	0.27
27	IN8Z	Extracellular domain of human HER2/therapeutic Fab	2.52	0.284
28	ISY6	Crystal structure of CD3 $\gamma\epsilon$ heterodimer/therapeutic Fab	2.1	0.255
29	IYY9	Extracellular domain of the epidermal growth factor receptor/neutralizing Fab	2.61	0.289

Our study is based on analysis of 29 Fab-antigen complex crystal structures (listed above) with resolution range 1.2-2.61 Å. All the Fabs in this study are different. This dataset compares favorably with the previous studies of Fab-antigen complex crystal structure data analyses (32,35,36). For example, Jackson et al. (1999) had studied 15 Fab-antigen complexes with resolution range 1.8-3.0 Å (32). The dataset of Lo Conte et al. (1999) contained 19 Fab-antigen complexes with resolution range 1.8-3.0 Å (35). The study of MacCallum et al. (1996) was based on 26 Fab-antigen complexes with 1.8-3.1 Å (36)

Contribution of APRs towards antigen recognition is measured by their contributions to the surface area of Fab buried upon complex formation. Amino acid residues from CDR-localized APRs contribute towards antigen binding in 28 out of 29 Fab-antigen complexes in our dataset. The percentage contributions are shown in Table II (column 5). On average,  $16.0\pm10.7\%$  (range: 0–42.7%) of buried surface area of Fabs can be attributed to APRs. The variation is large. For example, in case of the murine antibody Fab-protein A complex (PDB ID: 10SP), the variable domain contains only one APR which falls in H3 loop but does not participate in antigen binding. In the case of the VEFG-blocking Fab—Neuropilin-1 complex (PDB ID: 2QQN), the contribution of APRs is the highest (42.7%). APRs contribute more than 10% to buried surface area of Fab in 20 out of the 29 complexes (69%). Among commercial antibodies, with the exception of the neutralizing Fab-VEGF complex (PDB ID: 1BJ1), APRs account for greater than 15% of buried surface area of Fab. Fig. 3 shows, as an example, the structure of the IGG RU5 Fab-Von Willebrand factor complex (PDB ID: 1FE8). In this case, 29.8% of buried surface area in Fab is attributed to binding residues that belong to APRs. These observations indicate aggregation may be coupled with antigen binding function of antibodies.

The polar fractions of buried surface areas of Fabs and the share of APRs towards these fractions are shown in Fig. 4a and b, respectively. Polar surface area calculated in this study consists of the buried surface areas from both neutral polar and charged atoms. On average, the Fab part

1984 L.         -DiAMUTOTSHAAALAQVITATIAAGA	а						
17H3        DIMPERSIDENCE VITE SAGEDIMPERSIDENCE ON PREPROSECTIVE TRUE TO MANY DEPENDENCE VIEW AND THE INFORMATION OF THE ADDRESS AND ADD	1FE8 LC	-DIAMTOTTSSLSASLGOKVTIS	RASODIGNYLNWYOOKPI	OGTVRLLIY <mark>YTSRLHS</mark> GV	PSRFSGSGSGTDYSLTISNLESEDI.	ATYF <mark>CONGGTNPW</mark>	TFGGGTKLEVK-RADAA 112
<ul> <li>LIDD CILAURGESTALESPERATURE RANGE - FESETLAWY OPERATIVE NUMBER STATEMENT PROFESSORS OF THE ISSUE FRANK OF THE OUTCOME AND ALL INFORMATION AND ALL INF</li></ul>	1FNS LC						
LIVE_U         COUNTOBERSELASYORYTTTIALGAR         DIEYIANYORYKAKHOLIN ATLLANVPERSONSOTTUTUILISLEPENTY LIVENENT         TYPENTY LIVENENT         TYPENTY LIVENENTY           LIKI_C         COUNTOBERSELASYORYTTTIALGAR         STERMAN QUERQUERLANVERSONSOTTUTUILISLEPENTY LIVENENTY         TYPENTY LIVENENTY         TYPENTY	1H0D_LC	-DIVLTQSPASLAVSLGQRATIS	RASESVDNYGISFMSW <mark>FQQKP(</mark>	GQPPKLLIY <mark>AASNQGS</mark> GV	PARFSGSGSGTDFSLNIHPMEEDDT.	AMYF <mark>CQQSKEVPL</mark>	TFGAGTKLELK-RADAA 116
<ul> <li>HEB_LC - 010FTGERALGASUBTYTTETABLENITYELINVQUQUKSEQLUVNANTLEGUVESPESSAGESTOPSELTESUGDESEY UNDER</li></ul>							
11kb_cc         0100000000000000000000000000000000000							
<ul> <li>HUL_G. POILINGFAILUPTERSTELEVTERSTE</li></ul>	_						
1001_C0         >101M002005879VL00RVTTE MAGEDISTREMYOQUPCHANELLIGATELLOW PERSONSCRUTTETISSUPERFORMATING (MERSREPORTING INTER							
IUI_C         -DIMMOGRESLABVORDENTIT RAGCDISERSAMON (PERAFERS). INTRUELLO VIPSERGOSOTTUTINES, DEPERTITUE, DOBBEDIVENTITE RAGCENTRUM 1000000000000000000000000000000000000							
INE_C         D:OTTO: D:OTTO: BASE-ARVETYTI T: BASE FILIPIT.ANTO: DOCUMENT: FILIPINE DIFFERENCE PERFORMANCE PERFORMACE PERFORMANCE PERFORMACE PERFORMANCE PERFORMANCE PERFOR							
1YU_LC         KUILUTGERLIKAAPGEBYNEN         SABS							
12TL_C       -DUWTOBJEKEMESTSVORUSTITERAGBUSTANATYCKOROACHULTUVENET/DOGASCITUTTLISUC/DEDIATYTYCKORUTUT.BADAA 112         22NL_C       -DUKTOBJEKEMESTGANUTTITERAGQUENERATYCKOROACHULTUVENET/SOSGOSTITUTLISUC/DEDIATYTYCKORUTUR-HATAA 112         20NL_C       -DUKTOBJEKEMESTGANUTTITERAGQUENERATYCKOROACHULTUVENET/SOSGOSTITUTLISUC/DEDIATYTYCKORUTUR-HATAA 112         20NL_C       -DUKTOBJEKEMESTGANUTTITERAG							
<ul> <li>BASZ_LC - OLUTOSPSILSAS/UGRVITITAGAS</li></ul>							
2cm_LC         >DIOMTOSPETAS/EGRVTTITABLEGTUTULAN_COURGAPYLVIVOLUMEAD ISSUEDSOCIATITIESUCEDEPATY DOTINGLEPOSTLEIK-FUNA 112           2tm_LC         >DIVITOSPETAS/EGRVTTITABLEGTUTULAN_COURGAPYLVIVOLUMEAD ISSUEDSOCIATITIESUCEDEPATY COURSESSON WITHINGTON THE CONTROL NOT AND AND ADDITIESUCEDEPATY COURSESSON WITHINGTON THE CONTROL NOT ADDITIESUCEDEPATY COURSESSON WITHINGTON COURSESSON W							
2DB 1C - SPELTOPP - SYSTANCKTARITIGGUN LOESKUN YQUGYGAYUUVUDBORBG I JERFSGASGASTYSUIIJSKABDAUYUGYGUYUTUGPEN 113 2KK LC - DUWTOPPENTEUSOBERATLGVABE BYBLINY VQUGYGASTRULIYUKABCBASGASGASTYSUIJSKABDAUYUGYUNPENTEUSOTULIKRUNA 114 2001 LC - DUWTOPPENTEUSOBERATLGVABE BYBLINY VQUGYGASTRULIYUKABCBASGASGASTYSUIJSKABDAUYUGYUNPENTEUSOTULIKRUNA 112 2001 LC - DUWTOPPENTEUSOBERATLGVABE BYBLINY VQUGYGASTRULIYUAGTAGUYARFSGASGASGASTYSUIJSKABDAUYUGYUNPENTEUSOTULIKRUNA 112 2001 LC - DUWTOPPENTEUSOBUNTISTI TAGA YBBLINY VQUGYGASKULIYUBABTAGUYARFSGASGASTYTITISLQFEDATI'N UTLEF PFOQUTIVIKRUNA 112 2001 LC - DUWTOPPENTEUSOBUNTISTI TAGA YBBLINY VQUGYGASKULIYUBABTAGUYARFSGASGASTYTITISLQFEDATI'N UTLEF PFOQUTIVIKRUNA 112 2005 LC - DUWTOPSENTEUSOBUNTISTI TAGA BISNIGUAGUYAGUYARFSGASGASTYTILISLQFEDATI'N UTLEF PFOQUTIVIKRUNA 112 2005 LC - DUWTOPSENTEUSOBUNTISTI TAGA BISNIGUAGUYAGUYARFSGASGASTYTILISLQFEDATI'N UNDEF FROGUTIVIKRUNA 112 2006 LC - DUWTOPSENTEUSOBUNTISTI TAGA BISNIGUAGUYAGUYARFSGASGASTYTILISLQFEDATI'N UNDEF FROGUTIVIKRUNA 112 2007 LC - DUWTOPSENTEUSOBUNTISTI TAGA BISNIGUAGUYAGUYARFSGASGASTYTITISLQFEDATI'N UNDEF FROGUTIVIKRUNA 112 2007 LC - DUWTOPSENSALSANGUYUTITARAQ BISNIGUAGUYAGUYARFSGASGASTYTITISLQFEDATI'N UNDEF FROGUTIVIKRUNA 112 2007 LC - DUWTOPSENSALSANGUYUTITARAQ BUNYANYUQUKGHAAKULIY HISBLIGUYARFSGASGASTYTITISLQFEDATI'N UNDEF FROGUTIVIKRUNA 112 2007 LC - DUWTOPSENSALSANGUYUTITARAQ BUNYANYUQUKGHAAKULIY MARYUDAYBRYSGASGASTYTITISLAPEDATI'N UNDEF FROGUTIVIKRUNA 112 2007 LC - DUWTOPSENSALSANGUYUTITARAQ BUNYANYUQUKGHAAKULIY MARYUDAYBRYSGASGASTYTITISLANGUNTUN FROGUTIVIKRUNA 112 2007 LC - D'UVIGSENSALSANGUYUTITARAQ BUNYANYUQUKGHAAKULIY MARYUDAYBRYSGASGASTYTITISLANGUNTUN FROGUTIVIKRUNA 112 2007 LC - D'UVIGSENSALSANGUNYUTIANGUNGUKASUNUN UNDEFNILANGUNTUN UNDEFNILANGUNTUN FROGUTIVIKRUNA 112 2007 LC - D'UVIGSENSALSANGUNYUTINAYUQUKGHAAKULIY MARYUDAYBRYSGASGASTYTITISLANGUNTUN UNDEFNILANGUNUN TROGUTIVIKRUNA 112 2007 LC - D'UVIS							
<ul> <li>XIX_LC -DIVERGENTLES BASESUBJENT QOKPOGARRILLY ASTRUTY PERGENSIATE/TITISLG/BEDENTY QOKINGEN-EPFOGUNATE/TITISLG/BEDENTY QOKINGEN-EPFOGUNATE/TITISLG/BEDENTY QOKINGEN-EPFOGUNATE/TITISLG/BEDENTY QOKINGEN-EPFOGUNATE/TITISLG/BEDENTY QOKINGEN-EPFOGUNATE/TITISLG/BEDENTY QOKINGEN-EPFOGUNATE/TITISLG/BEDENTY QOKINGEN/ELIN-RADA 112</li> <li>ZON_LC -DIORTOSESILAS/SUDBUTTI RAGUSILMY QOKINGENK/LLY ASTRUTY DEFOGUNATE/TITISLG/BEDENTY QOKINGEN/ELIN-RADA 112</li> <li>ZVB LC -DIURTOSESILAS/SUDBUTTI RAGUSILMY QOKINGENK/LLY ASTRUTY DEFOGUNATI/TISLG/BEDENTY QOKINGEN/ELIN-RADA 112</li> <li>ZVD LC -DIURTOSESILAS/SUDBUTTI RAGSIGNILMY QOKINGENK/LLY ASTRUTY DEFOGUNATI/TISLG/BEDENTY QOKINGEN/ELIN-RADA 112</li> <li>ZDD LC -DIURTOSESILAS/DOWNTI BAGOSIGNILMY QOKINGEN/ELIN-RADA 113</li> <li>ZDD LC -DIURTOSESILAS/DUDITITI RAGDISTIMUM YQOKINGEN/ELIN-RADA 114</li> <li>ZDD LC -DIURTOSESILAS/DUDITITI RAGDISTIMUM YQOKINGEN/ELIN-RADA 112</li> <li>ZDD LC -DIURTOSESILAS/DUDITITI RAGSIGNILH/YAGENER/ENGON/LINATODES/DITISIS/DUTINES/LE/N-RADADISTIMUM YQOKINGEN/ELIN-RADA 112</li> <li>ZDD LC -DIURTOSESILAS/DAX/DUNAL/YAGOKINGEN/ELIN-RADADIVINGENCE/NINTER/LE/N-RADADIVINGENCE/NINTER/LE/N-RADADIVINGENCE/NINTER/LE/N-RADADIVINGENCE/NINTER/LE/N-RADADIVINGENCE/NINTER/LE/N-RADADIVINGENCE/NINTER/LE/N-RADADIVINGENCE/NINTER/LE/N-RADADIVINGENCE/NINTER/LE/N-RADADIVINGENCE/NINTER/LE/N-RADADIVINGENCE/NINTER/LE/N-RADADIVINGENCE/NINTER/LE/N-RADADIVINGENCE</li></ul>	2DD8 LC						
2001 L. STUMUTERLEVAGDAVUT I KAGOVIDENUM VOQKEGOSKILLYMASIRUVDEPTIGSKOTPTIISUPPERISTOVEELKINV QQFBED-TYGULGSEPPGUTKELKRADA 112 2001 L. SUUGYSSELSASUDEVITTERAG	2FD6_LC	-DIVLTQSPDITAASLGQKVTIT	SASSSVSYMHWYQQKS	GTSPKPWIF <mark>EISKLAS</mark> GV	PARFSGSGSGTSYSLTISSMEAEDA	AIYY <mark>C</mark> QQWNYPF	TFGGGTKLEIK-RADAA 110
<pre>200 Lc D-DIMTOSPSLASVODVITITAGAGPUSILAVOCKGARKLINGASEBASUVSRFSGGSGTDFTTISLOPEDATYY CUTGAST-FTFGGTNUEK-FTVA 112 2VB Lc D-DIMTOSPSLASVODVITIGAGDISNIGNLOVCKGARKLINGASEGASTASTISLOPEDATYY CUTGAST-FTFGGTNUEK-FTVA 112 2VB Lc D-DIMTOSPSMUSADVODVITIGAGDISNIGNLOVCKGASTKGLINGASTLINGASEBASUVSRFSGGSGDTFTISLOPEDATYY CUTGAST-FTFGGTNUEK-FTAADA 113 205 Lc D-DIMTOSPSMUSADVODVITIGAGBISNIGNLOVCKGASTKGLINGASGASTGTTISLOSEDPATYY CUTGASTFC-TTFGGTNUEK-FTAADA 113 205 Lc D-DIMTOSPSMUSADVODVITIGAGBISNIGNLOVCKGASTKGLINGASGGTSTTISLOSEDPATYY CUTGASTFC-TTFGGTNUEK-FTAADA 113 105 LC D-DIMTOSPSMUSADVODVITIGAGBISNIGNLOVCKGASTKGLINGASGGTSTTISLOSEDPATY CUTGASTFC-TTFGGTNUEK-FTAADA 113 105 LC D-DIMTOSPSMUSADVODVITIGAGBISNIGNLOVCKGASTKGLINGASGGTSTDTISLOSEDPATY CUTGASTFC-TTFGGTNUEK-FTAADA 113 105 LC D-DIMTOSPSMUSADVODVITIGAGBISNIGNLOVCKGASTKGLINGASGGTSTDTISLOSEDPATY CUTGASTFC-TTFGGTNUEK-FTAADA 113 105 LC D-DIMTOSPSMUSADVODVITIGAGBISNIGNLOVCKGASTKGLINGASGGTSTDTISLOSEDPATY 11400000000000000000000000000000000000</pre>	2NXY_LC	-DIVMTQSPATLSVSPGERATLS	<mark>RASESVSSDLAW</mark> YQQKP(	GQAPRLLIY <mark>GASTRAT</mark> GV	PARFSGSGSGAEFTLTISSLQSEDF.	AVYY <mark>C</mark> QQYNNWPPRY	TFGQGTRLEIK-RTVAA 114
<pre>sto_Lc =-DIMTOSPSILSASVODPVITTERAGCCUSSILVANTOQCHORGAPKULTYOSTULOVYERFSGGSGADVSILTISELOSEDPATYY CONTETTPOQCTVEILRTVAA 112 3D45_LC =-DIMTOSPSILSASVODPVITTERAGCGISSILUTIANTOQCHORGAPKULTYOTULOVYERFSGGSGADVSILTSELOSEDPATYW CURLEYTPOQCTVEILRTVAA 112 3D45_LC =-DIMTOSPSILSASVODPVITTERAGCCUSSILUTIANTOQCHORGAPKULTYERFSGGSGADVSILTSELOSEDPATYW CURLEYTPOQCTVEILRTVAA 112 1B1_LC =-DIMTOSPSILSASVODPVITTERAGCDISNILWY CORKIGARKULTYERFSGGSGADVSILTSELOSEDPATYW CURLEYTPOQCTVEILRTVAA 112 1B1_LC =-DIMTOSPSILSASVODPVITTERAGCDISNILWY CORKIGARKULTYERFSGGSGADVSILTSELOSEDPATYW CURLEYTPOQCTVEILRTVAA 112 1B1_LC =-DIMTOSPSILSASVODPVITTERAGCDISNILWY CORKIGARKULTYERFSGGSGADVFILTSELOPEDATYW CURLEY-RTVAA 112 1B1_LC =-DIMTOSPSILSASVODPVITTERAGC</pre>							
VDE_LC       -DILMTOSPESSINGUESTICTINGSO							
<ul> <li>abes_Lc - DIVMTOSPATLS THEORYSLS #RAG0 STBYLHHWYGKSHESPELLIKYAGGIS GISPEFSGSGSTDTI-LINNSTEEPKHYTY COURSEPT TPRGGTKLEIK - RADA 112</li> <li>abd_Lc - DIVMTOSPATLS THEORYSLS #RAG0 STBYLHWYGKSHESPELLIKYAGGIS GISPEFSGSGSGSDTTI-LISNSTEEPKHYTY COURSEPT TPRGGTKLEIK - RADA 112</li> <li>c - DIVMTOSPESLSAS GROWTIT #RAG0 DISHYLMY QCKPKAPKULITY TSSLUPE GISPEFSGSGSGTDTTITISSLOPED FAITY CUTTEP PERGGTKUEK - RTVA 112</li> <li>c - DIVMTOSPESLSAS GROWTIT #RAG0 DUWTAVAN QCKPKAPKULITY TSSLUPE GISPEFSGSGSGTDTTITISSLOPED FAITY CUTTEP PERGGTKUEK - RTVA 112</li> <li>c - DIVMTOSPESLSAS GROWTIT #RAG0 DUWTAVAN QCKPKAPKILI YNNAKGVYSPEFSGSGSGTDTTITISSLOPED FAITY CUTTEP PERGGTKUEK - RTVA 112</li> <li>c - DIVMTOSPESLSAS GROWTIT #RAG0 DUWTAVAN QCKPKAPKILI YNNAKGVYSPEFSGSGSGTDTTITISSLOPED FAITY CUTTEP PERGGTKUEK - RTVA 112</li> <li>c - DIVMTOSPESLSAS GROWTIT #RAG0 DUWTAVAN QCKPKAPKILI YNNAKGVYSPEFSGSGGGTDTTISIS WEED FAITY CUTTEP PERGGTKUEK - RTVA 112</li> <li>c - DIVMTOSPESLSAS GROWTIT #RAG0 DUWTAVAN QCKPKAPKILI YNNAKGVYSPEFSGSGGGTDTTISIS WEED TAYY CUTTEP PERGGTKUEK - RTVA 112</li> <li>c - DIVITOSPENSES BRAG0 SIGNIHH YQCRINGSPERG PAITHGGSGGTSSTSTISSET LITKDEYERHNSYT EATHKTSTS PIVKSFNR 211</li> <li>1 PTS LF PPSSEQLTSGGASVV FLANYPYKDINVKKIIGSERQNGVLIASTDQDSKDSTYSMSSTLTLTKDEYERHNSYT EATHKTSTS PIVKSFNR 211</li> <li>1 PTS LF PPSSEQLTSGGASVV FLANYPYKDINVKKIIGSERQNGVLIASTDQDSKDSTYSLSSTLTLTKDEYERHNSYT EATHKTSTS PIVKSFNR 210</li> <li>1 LOD LC PVTS IPPSSEQLTSGGASVV FLANYPYKDINVKKIIGSERQNGVLIASTDQDSKDSTYSLSSTLTLTKDEYERHNSYT EATHKTSTS PIVKSFNR 211</li> <li>1 LSS LC PVTS IPPSSEQLTSGGASVV FLANYPYKDINVKKIIGSERQNGVLIASTDQDSKDSTYSLSSTLTLTKDEYERHNSYT EATHKTSTS PIVKSFNR 210</li> <li>1 LMS LC PTVS IPPSSEQLTSGGASVV FLANYPYKDINVKKIIGSERQNGVLIASTDQDSKDSTYSLSSTLTLTKDEYERHNSYT EATHKTSTS PIVKSFNR 210</li> <li>1 LMS LC PTVS IPPSSEQLTSGGASVV FLANYPYKDINVKKKIIGSERQNGVLIASTDQDSKDSTYSLSSTLTLTKDEYERHNSYT EATHKTSTS PIV</li></ul>							
<pre>jbb_L c = brUngepartis/trees/sub_stage</pre>							
<ul> <li>IBJ_C - DIQMTOSPESLBASVDEVTITS ASQ DISMTUNEYQOKPGKAPKULTYTTSBLBGUPERPSGOGGOTPTLTISSUPEPATYTKOUTYVEKTPGOTTVUELK-RTVAA 112</li> <li>INZ LC - DIQMTOSPESLBASVDEVTITSAQ DUNTAVAMYQOKPGKAPKLLIYTBASTLG VDERPSGOGGOSTPTLSISUPEDIATYTKOUNTVEFTPGOTTVUELK-RTVAA 112</li> <li>INZ LC - DIQMTOSPESLBASVDEVTITSAQ DUNTAVAMYQOKPGKAPKLLIYTBASTLG VDERPSGOGGOSTPTLSISUPEDATYTYOUNTEFTPGOTTVUELK-RTVAA 112</li> <li>INZ LC - DIUTUOSPHILASASDEEKVITTSAASD SUTTNIHKYQOKTGSPELLIYTASSTLG VDERPSGOGGOSTPTLSISUPEDATYTYOUNTEFTPGOTTVUELK-RTVAA 112</li> <li>INZ LC - DILTOSPHILSVPGEVRSFERAQ SITTNIHKYQOKTGSPELLIYTASSTLG UPEPAFOSGOGGOTPTLSISUPEDATYTYOUNTEFTPGOTTVUELK-RTVAA 112</li> <li>INZ D 20 TAECD 30 40 50 50 70 80 59 586 100 110</li> <li>IPSSIPPSSEQLISGGASVV FLNNFYEKD INVKKKLDGGERQNOVLNSWTDODSKDSTYSMSSTLTLIKDEVERINSYT EATHKTSTSPIVKSFNRME 211</li> <li>IPSSIPPSSEQLISGGASVV FLNNFYEKD INVKKKLDGGERQNOVLNSWTDODSKDSTYSMSSTLTLIKDEVERINSYT EATHKTSTSPIVKSFNRME 214</li> <li>INDD_C PTVSIPPSSEQLISGGASVV FLNNFYEKD INVKKKLDGGERQNOVLDSKDSTYSMSSTLTLIKDEVERINSYT EATHKTSTSPIVKSFNRME 214</li> <li>IJSE_LC SPYFIPPSDEQLKSGTASVV LLNNFYFRBAKVQKKVDALQSCNSQESVTEQDSKDSTYSLSSTLTLKADYEKHKVYA EVTHQGLSSPVTKSFNRME 214</li> <li>ILKS_LC PTVSIPPSSEQLISGGASVV FLNNFYFKDINVKKKLDGERQNOVLDSKTOSTSMSSTLTLIKDEVERINSYT EATHKTSTSPIVKSFNRME 214</li> <li>ILKS_LC PTVSIPPSSEQLISGGASVV FLNNFYFKDINVKKKLDGERQNOVLDSKTODDSKDSTYSMSSTLTLIKDEVERINSYT EATHKTSTSPIVKSFNRME 214</li> <li>ILKS_LC PTVSIPPSSEQLISGGASVV FLNNFYFKDINVKKKLDGERQNOVLDSKTDODSKDSTYSMSSTLTLIKDEVERINSYT EATHKTSTSPIVKSFNRME 214</li> <li>ILKS_LC PTVSIPPSSEQLISGGASVV FLNNFYFKDINVKKKLDGERQNOVLDSKTDODSKDSTYSMSSTLTLIKDEVERINSYT EATHKTSTSPIVKSFNRME 214</li> <li>ILKS_LC PTVSIPPSSEQLISGGASVV FLNNFYFKDINVKKKLDGERQNOVLDSKTDODSKDSTYSMSSTLTLIKDEVERINSYT EATHKTSTSPIVKSFNRME 214</li> <li>ILVS_LC PTVSIPPSSEQLISGGASVV FLNNFYFKDINVKKKLDGERQNOVLDSKTDODSKDSTYSMSSTLTLIKDEVERINSYT EATHKTSTSPIVKSFNRME 214</li> <li>ILVS_LC PTVSIPPSSEQLISGGA</li></ul>							
<pre>ICRL/C -DIOMIGSESLASVORUTITE KABQWIDKILMYCOKPEKAPELLIYSEN (#VDERFSGSSGSGSGSGSGTPFPTISSLOPEDATYYGUILSEP-EFF0GGTVVSILK-FYVAA 112 ISYGL/C -OTVLOSPAINSASPORUTITE KABQWIDKILMYCOKSTEPKENTYDTSKLASGVDAPHRSGSGSTSYSLIISGKAEDATYYGOMSENEFFFGGTKLEIN-RADTA 111 IYY9_LC -OTVLOSPAINSASPORUTITE SABSSVSUMMYCOKSGTEPKENTYDTSKLASGVDAPHRSGSGGTSYSLIISGKAEDATYYGOMSENEFFFGGTKLEIN-RADTA 111 IYY9_LC -OTVLOSPAINSASPORUTITE SABSSVSUMMYCOKSGTEPKENTYDTSKLASGVDAPHRSGSGGTSYSLIISGKAEDATYYGOMSENEFFFGGTKLEIN-RADTA 111 IYY9_LC -OTVLOSPAINSASPORUTITE SABSSVSUMMYCOKSGTEPKENTYDTSKLASGVDAFHRSGSGGTSYSLIISGKAEDATYYGOMSENEFFFGGTKLEIN-RADTA 111 IYY9_LC -OTVLOSPAINSASPORUTITE SABSSVSUMMYCOKSGTEPKENTYDTSKLASGVDAFHRSGSGGTSYSLIISGKAEDATYYGOMSENEFFFGGTKLEIN-RADTA 111 IYY9_LC -OTVLOSPAINSASPORUTITE SAGASVVFENNSTERSTICHLYGONGUNNETFFGGTKLEIK-FKADTA IIFS-LC PTVSIFPPSSEQLITSGGASVVFENNYTEKKILGSERQNGVLINSWTDODSKDSTYSMSSTILTIKKDEYERINSYT EATHKTSTSPIVKSFNR.E- 214 IDSE_LC PTVSIFPPSSEQLISGGASVVFENNYTEKKINGSERQNGVLINSWTDODSKDSTYSISSTILTIKKDEYERINSYT EATHKTSTSPIVKSFNR.E- 214 IKS5_LC PTVSIFPPSSEQLISGGASVVFENNYFEKDINVKKKILGSERQNGVLINSWTDODSKDSTYSMSSTILTIKKDEYERINSYT EATHKTSTSPIVKSFNR.E- 214 IKS6_LC PTVSIFPPSSEQLISGGASVVFENNYKKKILGSERQNGVLINSWTDO</pre>							
<ul> <li>INRE_CC</li> <li>DIONTOSPESLASAVURVITIT HABC UVNTIVAM VOCKRAKPELLIY MASTI-SGUVERSGUSASTOPTITIS LOPEDATIV 00HTTPPTF00TTVEIL-RTVAA 112</li> <li>ISYE_LC</li> <li>DIULTOSPULSVSPGERVSFS</li> <li>RABCSIGTNIHN VOCRTAGSPERLIYASSGUSTSVLITSGURSLEDATVY 00NNMPTTFGATKLELK-RTVAA 112</li> <li>DIULTOSPULSVSPGERVSFS</li> <li>RABCSIGTNIHN VOCRTAGSPELLIKVASSGUSTSVLITSGURSLEDATVY 00NNMPTTFGATKLELK-RTVAA 112</li> <li>DIULTOSPULSVSPGERVSFS</li> <li>RABCSIGTNIHN VOCRTAGSPELLIKVASSGUSTSVLISSTLITLTKDE VERHNSVT EATHKTSTSPIVKSFNR</li></ul>							
<pre>1st6_C -0TVLTOSPATMSASPGEKYTMT SASESUSTAMINVOGKSTESPKENTYDTSKLASEVPARHPRSGSGTSTSTLITIGMBAEDAATYTYOOKSSNEFTFGSGTKLEIN-RADTA 111 1YT9_LC -DTLTOSPYLLSSPGEKYTSS RASGSUSTAMINVOGKSTESPKENTSDESTESTSTSTSTSUSSEDLATYTYOOKSSNEFTFGSGTKLEIN-RADTA 112 0 10 20 TAECDB 30 40 50 60 70 80 90 5AB6 100 110 1F88_LC PTVSIFPPSSEQLTSGGASVW FLNNFYPKDINVKKKLDGSERQNGVLNSWTDODSKDSTYSNSSTLITLTKDEYBERHNSYTGATHKITSTSPTVKSFNRNE- 214 1H0D_LC PTVSIFPPSSEQLTSGGASVW FLNNFYPKDINVKKLDGSERQNGVLNSWTDODSKDSTYSNSSTLITLKDEYBERHNSYTGATHKITSTSPTVKSFNRNE- 216 110D_LC PSVFIFPPSDEQLKSGTASVW LLNNFYPREAKVQKKUDALQSGNSQESVTEQDSKDSTYSLSSTLITLSKADYEKHKVXAEVTHQGLSSPTKSFNRNE- 214 1K55_LC PTVSIFPPSSEQLTSGGASVW FLNNFYPKDINVKKLDGSERQNGVLNSWTDODSKDSTYSNSSTLITLKDEYBERHNSYTGATHKTSTSPTVKSFNRNE- 214 1K55_LC PTVSIFPPSSEQLTSGGASVW FLNNFYPKDINVKKKLDGSERQNGVLNSWTDODSKDSTYSNSSTLITLKDEYBERHNSYTGATHKTSTSPTVKSFNRNE- 214 1K55_LC PTVSIFPPSSEQLTSGGASVW FLNNFYPKDINVKKLDGSERQNGVLNSWTDODSKDSTYSNSSTLITLKDEYBERHNSYTGATHKTSTSPTVKSFNRNE- 214 1K55_LC PTVSIFPPSSEQLTSGGASVW FLNNFYPKDINVKKLDGSERQNGVLNSWTDODSKDSTYSNSSTLITLKDEYBERHNSYTGATHKTSTSPTVKSFNRNE- 214 1K55_LC PTVSIFPPSSEQLTSGGASVW FLNNFYPKDINVKKLDGSERQNGVLNSWTDODSKDSTYSNSSTLITLKDEYBERHNSYTGATHKTSTSPTVKSFNRNE- 214 1K55_LC PTVSIFPPSSEQLTSGGASVW FLNNFYPKDINVKKLDGSERQNGVLNSWTDODSKDSTYSNSSTLITLTKDEYBERHNSYTGATHKTSTSPTVKSFNRNE- 214 1K54_LC PTVSIFPPSSEQLTSGGASVW FLNNFYPKDINVKKLDGSERQNGVLNSWTDODSKDSTYSNSSTLITLKDEYBERHNSYTGATHKTSTSPTVKSFNRNE- 214 200LC PTVSIFPPSSEQLTSGGA</pre>	_						
1YY9_LCDILLTQSPVILSVSPGENSG_SCH_STANCSIGTNINW YQQRTNGSPRLLIKYASESIGSIGTSTINSVSEEDIADYY CONNWED-TTPGGASYKFLANEYKAS112010207ABCD8 3040506070809050610010011FN_LCPTTSIFPSSEQLTSGGASYKFLANEYKDINYKWILDGSERQAVLINSWTDQDSKDSTYSMSSTLTITKDEYERINSYTEATHKTSTSPIVKSFNR2141HOD_LCPTVSIFPSSEQLTSGGASYKFLANEYKRINUNGSERQANGVLINSWTDQDSKDSTYSMSSTLTITKDEYERINSYTEATHKTSTSPIVKSFNR2161DD_LCPSVFIFPSSEQLXSGTASYKLLINNEYPERAKVQWKUNALQSGNSQESYTEDDSKDSTYSMSSTLTITKDEYERINSYTEATHKTSTSPIVKSFNR2141LKSLCPTVSIFPSSEQLTSGGASVKFLANEYPERAKVQWKUNALQSGNSQESYTEDDSKDSTYSLSSTLTISKADVEKKKVA EVTHQLSSPVTKSFNR2141LKSLCPTVSIFPSSEQLTSGGASVKFLANEYPERAKVQWKUNALQSGNSQESYTEDDSKDSTYSMSSTLTITKDEYERINSYTEATHKTSTSPIVKSFNR2141LKSLOSP_LCPTVSIFPSSEQLTSGGASVKFLANEYPENINVKWKIDGSERQAUSUNSWTDDDSKDSTYSMSSTLTITKDEYERINSYTEATHKTSTSPIVKSFNR2141LKSLOSP_LCPTVSIFPSSEQLTSGGASVKFLANEYPENINVKWKIDGSERQAUSUNSWTDDDSKDSTYSMSSTLTITKDEYERINSYTEATHKTSTSPIVKSFNR2141LKSLOSP_LCPTVSIFPSSEQLTSGGASVKFLANEYPENINVKWKIDGSERQAUSUNSWTDDDSKDSTYSMSSTLTITKDEYERINSYTEATHKTSTSPIVKSFNR2141LKSLOSP_LCPTVSIFPSSEQLTSGGASVKFLANEYPENINVKWKIDGSERQAUSUNSWTDDDSKDSTYSMSSTLTITKDEYERINSYTEATHKTSTSPIVKSFNR2121LKSLCPTVSIFPSSEQLTSGGASVKFLANEYPENINVKWKIDGSERQAUSUNSWTDDDSKDSTYSMSSTLTITKDEYERINSYTEATHKTSTSPIVKSFNR2141LVSLOSP_LCPT	_						
010207AECD8 304050607080905AB61001FE8_LCPTTSIFPPSEQLTSGGASVVFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEVERNNSYTEATHKTSTSPIVKSFNREF2111FNSLCPTVSIFPPSESQLTSGGASVVFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEVERNNSYTEATHKTSTSPIVKSFNREF2161HOD_LCPSVFIFPPSDEQLKSGTASVVFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEVERNNSYTEATHATSTSPIVKSFNREF2111JPSLOPSVFIFPPSDEQLKSGTASVVFLNNFYPEAKVQWKVDNALQGSNGSESVTEDDSKDSTYSLSSTLTLSKADVEKKKVAEVTHQLSSPVTKSFNREF2141LK5_LCPTVSIFPPSEQLXGGASVVFLNNFYPEAKVQWKVDNALQGSNGSESVTEDDSKDSTYSMSSTLTLTKDEVERNNSYTEATHKTSTSPIVKSFNREF2141LK6_LCPTVSIFPPSEQLXGGASVVFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEVERNNSYTEATHKTSTSPIVKSFNREF2141LW2_LCPTVSIFPPSEDQLKSGTASGASVVFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEVERNNSYTEATHKTSTSPIVKSFNREF2141UJ3_LCPSVFIFPPSEDQLKSGTASGASVVFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEVERNNSYTEATHKTSTSPIVKSFNREF2141UJ3_LCPTVSIFPPSSEQLTSGGASVVFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEVERNNSYTEATHKTSTSPIVKSFNREF2141QV_LCPTVSIFPPSSEQLTSGGASVVFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEVERNNSYTEATHKTSTSPIVKSFNREF2141QV_LCPTVSIFPPSSEQLTSGGASVVFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEVERNNSYTEATHKTSTSPIVKSFNREF2141QV_LCPTVSIFPPSSEQLTSGGASVVFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYS							
<pre>1FE8_LC PTTSIFPPSSQLTSGGASVW PLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERHNSYTGEATHKTSTSPIVKSPNRFE</pre>		* * *	***** * *	*	* * *	* ****	* *
<pre>IND_LC PTVSIFPPSSEQLTSGGASVV FLNNFYPKDINVKWKLDGSERQMGVLNSWTDQDSKDSTYSMSSTLTLTKDEVERHNSYT EATHKTSTSFIVKSFNRME 214 IHOD_LC PSVFIFPPSDEQLKSGTASVV LLNNFYPKEAKVQWKUDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADVEKKKYA EVTHQGLSSPVTKSFNRME 211 IJFS_LC PSVFIFPSDEQLKSGTASVV LLNNFYPKEAKVQWKUDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADVEKKKYA EVTHQGLSSPVTKSFNRME 214 IKA5_LC PTVSIFPSSEQLTSGGASVV FLNNFYPKDINVKWKLDGSERQMGVLNSWTDQDSKDSTYSMSSTLTLTKDEVERHNSYT EATHKTSTSFIVKSFNRME 210 IMLC_LC PTVSIFPSSEQLTSGGASVV FLNNFYPKDINVKWKLDGSERQMGVLNSWTDQDSKDSTYSMSSTLTLTKDEVERHNSYT EATHKTSTSFIVKSFNRME 214 IUJ3_LC PTVSIFPSSEQLTSGGASVV FLNNFYPKDINVKWKLDGSERQMGVLNSWTDQDSKDSTYSMSSTLTLTKDEVERHNSYT EATHKTSTSFIVKSFNRME 214 IUZY_L PTVSIFPSSEQLTSGGASVV FLNNFYPKDINVKWKLDGSERQMGVLNSWTDQDSKDSTYSMSSTLTLTKDEVERHNSYT EATHKTSTSFIVKSFNRME 214 IZYLC PTVSIFPSSEQLTSGGASVV FLNNFYPKDINVKWKLDGSERQMGVLNSWTDQDSKDSTYSMSSTLTLTKDEVERHNSYT EATHKTSTSFIVKSFNRME 218 ZDME_L PTVSIFPSSEQLTSGGASVV FLNNFYPKDINVKWKLDGSERQMGVLNSWTDQDSKDSTYSMSSTLTLTKDEVERHNSYT EATHKTSTSFIVKSFNRME 214 20KM_L PSVFIFPSSEQLTSGGASVV FLNNFYPKDINVKWKLDGSERQMGVLNSWTDQDSKDSTYSMSSTLTLTKDEVERHNSYT EATHKTSTSFIVKSFNRME 214 20KL PSVFIFPSSEQLTSGGASVV FLNNFYPKDINVKKULGSERQMGVLNSTDQDSKDSTYSMSSTLTLTKDEVERHNSYT</pre>	1000 10		_				
1H0D_LCPTVSIFPPSSEQLTSGGASVV FLNNFYPKDINVRWKIDGSERONGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERHNSYT EATHAASPIVKSFNRME							
11QD_LC PSVFIFPPSDEQLKSGTASVV LLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSPNR 211 1JF5_LC PTVSIFPPSSEQLXSGTASVV LLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE 214 1KK5_LC PTVSIFPPSSEQLTSGGASVV FLNNFYPKDISVKWKLDGSERQNGVLNSWTQDSKDSTYSMSSTLLTKADYERHKVYACEVTHQGLSSPVTKSFNRGE 214 1KLC_LC PTVSIFPPSSEQLTSGGASVV FLNNFYPKDISVKWKLDGSERQNGVLNSWTQDSKDSTYSMSSTLTLTKDEYERHNSTTCEATHKTSTSPIVKSFNRNE 214 103P_LC PTVSIFPPSSEQLTSGGASVV FLNNFYPKDINVKWKLDGSERQNGVLNSWTQDSKDSTYSMSSTLTLTKDEYERHNSTTCEATHKTSTSPIVKSFNRNE 214 103P_LC PTVSIFPPSSEQLTSGGASVV FLNNFYPKDINVKWKLDGSERQNGVLNSWTQDSKDSTYSMSSTLTLTKDEYERHNSTTCEATHKTSTSPIVKSFNRNE 214 103J_LC PTVSIFPPSSEQLTSGGASVV FLNNFYPKDINVKWKLDGSERQNGVLNSWTQDSKDSTYSMSSTLTLTKDEYERHNSTTCEATHKTSTSPIVKSFNRNE 214 104L_L PTVSIFPPSSEQLTSGGASVV FLNNFYPKDINVKWKLDGSERQNGVLNSWTQDSKDSTYSMSSTLTLTKDEYERHNSTTCEATHKTSTSPIVKSFNRNE 211 127L_L PTVSIFPPSSEQLTSGGASVV FLNNFYPKDINVKWKLDGSERQNGVLNSWTQDSKDSTYSMSSTLTLTKDEYERHNSTTCEATHKTSTSPIVKSFNRNE 212 222K_L PTVSIFPPSSEQLTSGGASVV FLNNFYPKDINVKWKLDGSERQNGVLNSWTQDSKDSTYSMSSTLTLTKDEYERHNSTTCEATHKTSTSPIVKSFNRNE 213 200R_LC PTVSIFPPSSEQLTSGGASVV FLNNFYPKDINVKWKLDGSERQNGVLNSWTQDSKDSTYSMSSTLTLTKDEYERHNSTTCEATHKTSTSPIVKSFNRNE 214 202K_LC PSVFIFPPSSEQLTSGGASVV FLNNFYPKDINVKWKLDGSERQNGVLNSWTQDSKDSTYSMSSTLTLTKDEYERHNSTTCEATHKTSTSPIVKSFNRNE 214 202K_LC PTVSIFPPSSEQLSGGASVV FLNNFYPKDINVKWKLDGSERQNGVLNSWTQDSKDSTYSMSSTLTLTKDEYERHNSTTCEATHKTSTSPIVKSFNRGE 214 202K_LC PSVFIFPPSDEQLKSGTASVV LLNNFYPEEAKVQWKVDNALQGSNQESVTEQDSKDSTYSMSSTLTLTKDEYERHNSTTCEATHKTSTSPIVKSFNRGE 214 202K_LC PSVFIFPPSDEQLKSGTASVV LLNNFYPEEAKVQWKVDNALQGSNQESVTEQDSKDSTYSMSSTLTLTKDEYERHNSTTCEATHKTSTSPIVKSFNRGE 214 202K_LC PSVFIFPPSDEQLKSGTASVV LLNNFYPEEAKVQWKVDNALQGSNQESVTEQDSKDSTYSMSSTLTLTKDEYERHNSTTCEATHKTSTSPIVKSFNRGE 214 202K_LC PSVFIFPPSDEQLKSGTASVV LLNNFYPEEAKVQWKVDNALQGSNQESVTEQDSKDSTYSMSSTLTLTKDEYERHNSTTCEATHKTSTSPIVKSFNRGE 214 202K_LC PSVFIFPPSDEQLKSGTASVV LLNNFYPEEAKVQWKVDNALQGSNQESVTE							
1JPS_LC PSVFIPPSBQLKSGTASVV LLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAGEVTHQGLSSPVTKSFNRGE 214 1KB5_LC PTVSIFPPSBCQLTSGGASVV FLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYEHNSYTEATHKTSTSPIVKSFNRNE 210 1MLC_LC PTVSIFPPSBCQLTSGGASVV FLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTTKDEYEHNSYTEATHKTSTSPIVKSFNRME 214 10SP_LC PTVSIFPPSBCQLTSGGASVV FLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTTKDEYEHNSYTEATHKTSTSPIVKSFNRME 214 10J3_LC PSVFIFPPSDCQLKSGTASVV FLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTTKDEYEHNSYTEATHKTSTSPIVKSFNRME 214 11QV_LC PTVSIFPSSCQLTSGGASVV FLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTTKDEYEHNSYTEATHKTSTSPIVKSFNRME 214 12TX_LC PTVSIFPSSCQLTSGGASVV FLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTTKDEYEHNSYTEATHKTSTSPIVKSFNRME 214 12QL_C PTVSIFPPSSCQLTSGGASVV FLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTTKDEYEHNSYTEATHKTSTSPIVKSFNRME 212 22XL_C PTVSIFPPSSCQLTSGGASVV FLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTTKDEYEHNSYTEATHKTSTSPIVKSFNRME 213 20ML_C PSVFIFPSSCQLTSGGASVV FLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTTKDEYEHNSYTEATHKTSTSPIVKSFNRME 213 20LD _LC PTVSIFPPSSCQLTSGGASVV FLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYEHNSYTEATHKTSTSPIVKSFNRME 214 20NL_C PSVFIFPPSSCQLTSGGASVV FLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYEHNSYTEATHKTSTSPIVKSFNRME 214 20NL_LC PSVFIFPPSSCQLTSGGASVV FLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYEHNSYTEATHKTSTSPIVKSFNRME 214 20NL_LC PSVFIFPPSSCQLTSGGASVV FLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYEHNSYTEATHKTSTSPIVKSFNRME 214 20NL_LC PSVFIFPSSCQLTSGGASVV FLNNFYPKDINVKWKIDGSERQNGVSNSDDSKDSTYSMSSTLTLTKDEYEHNSYTEATHKTSTSPIVKSFNRME 214 20NL_LC PSVFIFPSDEQLKSGTASVV LLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSMSSTLTLTKDEYEHNSYTEATHKTSTSPIVKSFNRME 214 20NL_LC PSVFIFPSDEQLKSGTASVV LLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSMSSTLTLSKADYEKKVXAEEVTHQGLSSPVTKSFNRME 214 20NL_LC PSVFIFPSDEQLKSGTASVV LLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSMSSTLTLSKADYEKKVXAEEVTHQ				GSEKÖNGATUSMIDÖT	JSKUSTISMSSTLTLTKUEIERI		
1KB5_LCPTVSIFPPSSEQLTSGGASVV.FLNNFYPKDINVKWKIDGSERQNGVLNSWTQDSKDSTYSMSSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRRE				173 T O C C 17 C C C C C C C C C C C C C C C C			
<pre>1LK3_LC PTVSIFPPSTEQLATGGASVV LMNNFYPRDISVKWKIDGTERRDGVLDSVTDQDSKDSTYSMSSTLSLTKADYESHNLYT EVVHKTSSSPVVKSFNR</pre>	1JPS_LC				OSKDSTYSLSSTLTLSKADYEK	ikvya <mark>c</mark> evthqgls	SPVTKSFNR 211
<pre>NMLC_LC PTVSIFPPSSeQLTSGGASVVFLINNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERNNSYTEATHKTSTSPIVKSFNRNEE 214 losp_LC PTVSIFPPSSeQLTSGGASVVFLINNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERNNSYTEATHKTSTSPIVKSFNRNEE 214 lwEJ_LC PTVSIFPPSSEQLTSGGASVVFLINNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERNNSYTEATHKTSTSPIVKSFNRNEE 214 lvGV_LC PTVSIFPPSSEQLTSGGASVVFLINNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERNNSYTEATHKTSTSPIVKSFNRNEE 214 lvGV_LC PTVSIFPPSSEQLTSGGASVVFLINNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERNNSYTEATHKTSTSPIVKSFNRNEE 212 lzTX_LC PTVSIFPPSSEQLTSGGASVVFLINNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERNNSYTEATHKTSTSPIVKSFNRNEE 212 l2CXR_LC PSVFIFPSSEQLTSGGASVVFLINNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERNNSYTEATHKTSTSPIVKSFNRNEE 212 l2CXR_LC PSVFIFPSSEQLTSGGASVVFLINNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERNNSYTEATHKTSTSPIVKSFNRNEE 214 lvGV_L PSVFIFPSSEQLTSGGASVVFLINNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERNNSYTEATHKTSTSPIVKSFNRNEE 214 lvGN_LC PSVFIFPSSEQLTSGGASVVFLINNFYPKDINVKKKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERNNSYTEATHKTSTSPIVKSFNRNEE 214 lvGN_LC PSVFIFPSSEQLTSGGASVVFLINNFYPKDINVKKKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERNSYTEATHKTSTSPIVKSFNRNEE 214 lvGN_LC PSVFIFPSSEQLTSGGASVVFLINNFYPKDINVKKKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERNSYTEATHKTSTSPIVKSFNRNEE 214 lvGN_LC PSVFIFPSSEQLTSGGASVVFLINNFYPKDINVKKKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERNSYTEATHKTSTSPIVKSFNREE 214 lvGN_LC PSVFIFPSSEQLTSGGASVVFLINNFYPKDINVKKKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERNSYTEATHKTSTSPIVKSFNREE 214 lvGL PSVFIFPSSEQLTSGGASVVFLINNFYPKDINVKKKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLSKADYEKHVVACEVTHQGLSSPVTKSFNREE 214 lvGL PSVFIFPSSEQLTSGGASVVFLINNFYPKDINVKKKIDGSERQNGVLDSVFLQDSKDSTYSMSSTLTLSKADYEKHVVACEVTHQGLSSPVTKSFNREE 214 lvGL PSVFIFPSDEQLKSGTASVVFLINNFYPKDINVKKKIDGSERQNGVLDSVFLQDSKDSTYSMSSTLTLSKADYEKHVVACEVTHQGLSSPVTKSFNREE 214 lvGL PSVFIFPSDEQLKSGTASVVFLINNFYPKDINVKKKIDGSERQNGVEDQDSKDSTYSMSSTLTLSKADYEKHVVA</pre>		P <mark>SVFIFP</mark> PSDEQLKSGTA <mark>S</mark> V	V <mark>C</mark> LLNNFYPREAKVQWKVD	NALQSGNSQESVTEQI	DSKDSTYSLSSTLTLSKADYEKI DSKDSTYSLSSTLTLSKADYEKI	IKVYA <mark>C</mark> EVTHQGLS IKVYA <mark>C</mark> EVTHQGLS	SPVTKSFNR 211 SPVTKSFNRGE <mark>C</mark> 214
10SP_LCPTVSIFPPSSEQLTSGGASVVFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLKKDEYERHNSYTGEATHKTSTSPIVKSFNRNEC2141UJ3_LCPSVFIFPPSDEQLKSGTASVVFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLKKDEYERHNSYTGEATHKTSTSPIVKSFNRNEC2141VQV_LCPTVSIFPPSSEQLTSGGASVVFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLKDEYERHNSYTGEATHKTSTSPIVKSFNRN2111ZTX_LCPTVSIFPPSSEQLTSGGASVVFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLKDEYERHNSYTGEATHKTSTSPIVKSFNRN212222_LCPTVSIFPPSSEQLTSGGASVVFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLKDEYERHNSYTGEATHKTSTSPIVKSFNRN2132CMR_LCPSVFIFPPSDEQLKSGTASVVFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLKDEYERHNSYTGEATHKTSTSPIVKSFNRNEC2182DD8_LCPTVSIFPPSSEEFQANKATLVLISDFYPGAVTVAWKADGSEVKAGVETTKPSKQS-NNKYAASSYLSLTPEQWKSHRSYSQVTHGGSTVEKTVATEGCS2142NXY_LCPSVFIFPPSSEQLTSGGASVVFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSLSSTLTLKADYEKHKVYAGEVTHQGLSSPVTKSFNRNEC2142Q0N_LCPTVSIFPPSSEQLTSGGASVVFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSLSSTLTLKADYEKHKVYAGEVTHQGLSSPVTKSFNRGEC2142Q0N_LCPSVFIFPPSDEQLKSGTASVVLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAGEVTHQGLSSPVTKSFNRGEC214200L_LCPSVFIFPPSDEQLKSGTASVVLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAGEVTHQGLSSPVTKSFNRGEC214200L_LCPSVFIFPPSDEQLKSGTASVVLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAGEVTHQGLSSPVTKSFNRGEC214200L_LCPSVFIFPPSDEQLKSGTASVVLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAGEVTHQGLSSPVTKSFNRGEC<		P <mark>SVFIFP</mark> PSDEQLKSGTA <mark>SV</mark> PTVSIFPPSSEQLTSGGA <mark>SV</mark>	V <mark>C</mark> LLNNFYPREAKVQWKVD V <mark>C</mark> FLNNFYPKDINVKWKID	NALQSGNSQESVTEQI GSERQNGVLNSWTDQI	DSKDSTYSLSSTLTLSKADYEKI DSKDSTYSLSSTLTLSKADYEKI DSKDSTYSMSSTLTLTKDEYERI	IKVYA <mark>C</mark> EVTHQGLS IKVYACEVTHQGLS INSYT <mark>C</mark> EATHKTSI	SSPVTKSFNR 211 SSPVTKSFNRGEC 214 SSPIVKSFNRNEC 214
1UJ3_LCPSVFIFPPSDEQLKSGTASVVLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVAGEVTHQGLSSPVTKSFNRGECT2151WBJ_LCPTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERNNSYT EATHKTSTSPIVKSFNRNEC-2141YQV_LCPTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERNNSYT EATHKTSTSPIVKSFNRN2111ZTX_LCPTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERNNSYT EATHKTSTSPIVKSFNRN2122B2X_LCPTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLKDEYERNNSYT EATHKTSTSPIVKSFNRNEC2132CMR_LCPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDDALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAEVTHQGLSSPVTKS2D8LCPTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLKDEYERNNSYT EATHKTSTSPIVKSFNRNEC	1LK3_LC	P <mark>SVFIFP</mark> PSDEQLKSGTA <mark>SV</mark> PTVSIFPPSSEQLTSGGA <mark>SV</mark> PTVSIFPPSTEQLATGGASV	V <mark>C</mark> LLNNFYPREAKVQWKVC VCFLNNFYPKDINVKWKIC V <mark>C</mark> LMNNFYPRDISVKWKIC	NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GTERRDGVLDSVTDQI	DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLSLTKADYES	IKVYA <mark>C</mark> EVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTSI INLYT <mark>C</mark> EVVHKTSS	SSPVTKSFNR 211 SSPVTKSFNRGEC 214 SSPIVKSFNRNEC 214 SSPVVKSFNR 210
<pre>1WEJ_LC PTVSIFPPSSEQLTSGGASVV FLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRNEG- 214 1YQV_LC PTVSIFPPSSEQLTSGGASVV FLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRN 211 1ZTX_LC PTVSIFPPSSEQLTSGGASVV FLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRN 212 2D2LK_LC PTVSIFPPSSEQLTSGGASVV FLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRNEG 213 2CMR_LC PSVFIFPPSDEQLKSGTASVV LLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKS</pre>	1LK3_LC	P <mark>SVFIFP</mark> PSDEQLKSGTA <mark>SV</mark> PTVSIFPPSSEQLTSGGA <mark>SV</mark> PTVSIFPPSTEQLATGGASV	V <mark>C</mark> LLNNFYPREAKVQWKVC VCFLNNFYPKDINVKWKIC V <mark>C</mark> LMNNFYPRDISVKWKIC	NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GTERRDGVLDSVTDQI	DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLSLTKADYES	IKVYA <mark>C</mark> EVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTSI INLYT <mark>C</mark> EVVHKTSS	SSPVTKSFNR 211 SSPVTKSFNRGEC 214 SSPIVKSFNRNEC 214 SSPVVKSFNR 210
1YQV_LCPTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERHNSYTCATHKTSTSPIVKSFNRN	1LK3_LC 1MLC_LC	PSVF1FPPSDEQLKSGTASV PTVS1FPPSSEQLTSGGASV PTVS1FPPSTEQLATGGASV PTVS1FPPSSEQLTSGGASV	V <mark>CLLNNFYPREAKVQWKVD</mark> VCFLNNFYPKDINVKWKID VCLMNNFYPRDISVKWKID V <mark>CFLNNFYPKDINVKWKID</mark>	NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GTERRDGVLDSVTDQI GSERQNGVLNSWTDQI	DSKDSTYSLSSTLTLSKADYEKI DSKDSTYSLSSTLTLSKADYEKI DSKDSTYSMSSTLTLTKDEYERI DSKDSTYSMSSTLSLTKADYESI DSKDSTYSMSSTLTLTKDEYERI	IKVYA <mark>C</mark> EVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTST INLYTCEVVHKTSS INSYT <mark>C</mark> EATHKTST	SSPVTKSFNR 211 SSPVTKSFNRGEC 214 SSPIVKSFNRNEC 214 SSPVVKSFNR 210 SSPIVKSFNRNEC 214
1ZTX_LC PTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERHNSYTCATHKTSTSPIVKSFNRNE	1LK3_LC 1MLC_LC 10SP_LC	PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSTEQLATGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV	VELLNNFYPREAKVQWKVD VEFLNNFYPKDINVKWKID VELMNNFYPRDISVKWKID VEFLNNFYPKDINVKWKID VEFLNNFYPKDINVKWKID	NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GTERRDGVLDSVTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI	DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSMSSTLTLTKADYER DSKDSTYSMSSTLSLTKADYES DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER	IKVYACEVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTST INLYTCEVVHKTSS INSYTCEATHKTST INSYTCEATHKTST	SPVTKSFNR211 SPVTKSFNRGE214 SPVTKSFNRNEC214 SPVVKSFNR210 SPVVKSFNRNEC214 SPIVKSFNRNEC214
1ZTX_LC PTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERHNSYTCATHKTSTSPIVKSFNRNE	1LK3_LC 1MLC_LC 10SP_LC 1UJ3_LC	PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PSVFIFPPSDEQLKSGTASV	V LLNNFYPREAKVQWKVD V FLNNFYPRDINVKWKID V FLNNFYPRDISVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID LLNNFYPREAKVQWKVD	NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GTERRDGVLDSVTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI	DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLSLTKADYES DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSLSSTLTLSKADYEK	IKVYA <mark>C</mark> EVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTST INLYTCEVVHKTSS INSYTCEATHKTST INSYTCEATHKTST IKVYACEVTHQGLS	SPVTKSFNR         211           SPVTKSFNRGE         214           SPVTKSFNRNE         214           SPVVKSFNRNE         214           SPVVKSFNR         214           SPVKSFNRE         214           SPVKSFNRE         214           SPVKSFNRE
2CMR_LCPSVFIFPPSDEQLKSGTASVVLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAGEVTHQGLSSPVTKS	1LK3_LC 1MLC_LC 1OSP_LC 1UJ3_LC 1WEJ_LC	PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQLTSGGASV	V LLNNFYPREAKVQWKVD V FLNNFYPRDINVKWKID V LMNNFYPRDISVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V LLNNFYPREAKVQWKVD V FLNNFYPRDINVKWKID	NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GTERRDGVLDSVTDQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI	DSKDSTYSLSSTLTLSKADYEKH DSKDSTYSLSSTLTLSKADYEKH DSKDSTYSMSSTLSLTKDEYERH DSKDSTYSMSSTLSLTKADYESH DSKDSTYSMSSTLTLTKDEYERH DSKDSTYSMSSTLTLTKDEYERH DSKDSTYSLSSTLTLSKADYEKH DSKDSTYSMSSTLTLTKDEYERH	IKVYACEVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTST INLYTCEVHKTST INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST INSYTCEATHKTST	SSPVTKSFNR         211           SSPVTKSFNRGE         214           SSPVKSFNRE
2CMR_LCPSVFIFPPSDEQLKSGTASVVLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAGEVTHQGLSSPVTKS	1LK3_LC 1MLC_LC 1OSP_LC 1UJ3_LC 1WEJ_LC 1YQV_LC	PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQLTSGGASV	V LLNNFYPREAKVQWKVD V FLNNFYPRDINVKWKID V LMNNFYPRDISVKWKID V FLNNFYPRDINVKWKID V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V FLNNFYPRDINVKWKID	NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GTERRDGVLDSVTDQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI	DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSMSSTLTLTKKDYER DSKDSTYSMSSTLSLTKADYES DSKDSTYSMSSTLSLTKADYES DSKDSTYSMSSTLTLTKDEYER DSKDSTYSLSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER	IKVYACEVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST IKVYACEVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTST INSYTCEATHKTST	SPVTKSFNR
2DD8_LC PTVTLFPPSSEEFQANKATLVCLISDFYPGAVTVAWKADGSPVKAGVETTKPSKQS-NNKYAASSYLSLTPEQWKSHRSYSC VTHEGSTVEKTVAPTEGS-213 2FD6_LC PTVSIFPPSSEQLTSGGASVVCFLNNFYPRDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERHNSYTCATHKTSTSPIVKSFNRNEAKA 214 2NXY_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKKKVACEVTHQGLSSPVTKSFNREG214 2Q0N_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKKKVACEVTHQGLSSPVTKSFNRGE214 2R0L_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKKKVACEVTHQGLSSPVTKSFNRGE214 2VDR_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKKKVACEVTHQGLSSPVTKSFNRGE214 3D85_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLKADYEKKKVACEVTHQGLSSPVTKSFNRGE214 3D85_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKKKVACEVTHQGLSSPVTKSFNRGE214 3D85_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKKVACEVTHQGLSSPVTKSFNRGE214 3D85_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKKVACEVTHQGLSSPVTKSFNRGE214 3D85_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKKVACEVTHQGLSSPVTKSFNRGE214 3D85_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKKVACEVTHQGLSSPVTKSFNRGE214 3D85_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKKVACEVTHQGLSSPVTKSFNRGE214 3D85_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKKVACEVTHQGLSSPVTKSFNRGE214 1S82_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKKVACEVTHQGLSSPVTKSFNRGE214 1S86_LC PTVSIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKKVACEVTHQGLSSPVTKSFNRGE214 1S96_LC PTVSIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKKVACEVTHQGLSSPVTKSFNRGE213 1YY9_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKKVACEVTHQGLSSPVTKSFNRGE213	1LK3_LC 1MLC_LC 1OSP_LC 1UJ3_LC 1WEJ_LC 1YQV_LC 1ZTX_LC	PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV	V LLNNFYPREAKVQWKVD V FLNNFYPRDISVKWKID V FLNNFYPRDISVKWKID V FLNNFYPRDINVKWKID V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID	NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI	DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSMSSTLTLTKKDYER DSKDSTYSMSSTLSLTKADYES DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSLSSTLTLSKADYEK DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER	IKVYACEVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST	SPVTKSFNR
2FD6_LC PTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRNEÅKA 214 2NXY_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSMSSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRBE 214 2Q&B_LC PTVSIFPPSSEQLTSGGASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSMSSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRBE 214 2QQN_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRBE 214 2NDR_LC PTVSIFPPSSEQLTSGGASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRBE 214 3D85_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRBE 214 3D85_LC PTVSIFPPSSEQLTSGGASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRBE 214 3D85_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRBE 214 3D85_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRBE 214 3D85_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRBE 214 3D85_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRBE 214 3D85_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRBE 214 3D85_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRBE 214 1N82_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRBE 214 1SY6_LC PTVSIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRBE 213 1YY9_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGA 213 1YY9_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGA 213	1LK3_LC 1MLC_LC 10SP_LC 1UJ3_LC 1WEJ_LC 1YQV_LC 1ZTX_LC 2B2X_LC	PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV	V LLNNFYPREAKVQWKVD V FLNNFYPRDISVKWKID V FLNNFYPRDISVKWKID V FLNNFYPRDINVKWKID V LLNNFYPRBAKVQWKVD V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID	NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLDSVTDQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI	DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLSLTKADYES DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSLSSTLTLSKADYEK DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER	IKVYACEVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST	SPVTKSFNR       211         SPVTKSFNRBE       214         SPVKSFNRBE       210         SPVKSFNRBE       214         SPVKSFNRNE       211         SPIVKSFNRN       212         SPIVKSFNRN       212         SPIVKSFNRNE       212         SPIVKSFNRNE       213
2NXY_LC PSVFIFPPSDEQLKSGTASVVLLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAGEVTHQGLSSPVTKSFNRGE 214 2Q8B_LC PTVSIFPPSSEQLTSGGASVVCFLNNFYPRDINVKWKUDGSERQNGVLNSWTDQDSKDSTYSLSSTLTLTKDEYERINSYTEATHKTSTSPIVKSFNRGE 214 2Q0N_LC PSVFIFPPSDEQLKSGTASVVLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAGEVTHQGLSSPVTKSFNRGE 214 2VDR_LC PTVSIFPPSSEQLTSGGASVVCFLNNFYPRDINVKWKUDGSERQNGVLNSWTDQDSKDSTYSLSSTLTLSKADYEKHKVYAGEVTHQGLSSPVTKSFNRGE 214 3D85_LC PTVSIFPPSSEQLTSGGASVVCFLNNFYPRDINVKWKUDGSERQNGVLNSWTDQDSKDSTYSLSSTLTLKDEYERINSYTEATHKTSTSPIVKSFNRGE 214 3D9A_LC PTVSIFPPSSEQLTSGGASVVCFLNNFYPRDINVKWKUDGSERQNGVLNSWTDQDSKDSTYSLSSTLTLKDEYERINSYTEATHKTSTSPIVKSFNRGE 214 3D9A_LC PTVSIFPPSSEQLTSGGASVVCFLNNFYPRDINVKWKUDGSERQNGVLNSWTDQDSKDSTYSLSSTLTLKNEYERINSYTEATHKTSTSPIVKSFNRGE 214 3D9A_LC PTVSIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLKNEYERINSYTEATHKTSTSPIVKSFNRGE 214 1BJ1_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAGEVTHQGLSSPVTKSFNRGE 211 1N8Z_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAGEVTHQGLSSPVTKSFNRGE 211 1N8Z_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAGEVTHQGLSSPVTKSFNRGE 214 1SY6_LC PTVSIFPPSSEQLTSGGASVVCFLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAGEVTHQGLSSPVTKSFNRGE 211 1N8Z_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAGEVTHGGLSSPVTKSFNRGE 213 1YY9_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAGEVTHGGLSSPVTKSFNRGE 213 1YY9_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAGEVTHQGLSSPVTKSFNRGA 213	1LK3_LC 1MLC_LC 1OSP_LC 1UJ3_LC 1WEJ_LC 1YQV_LC 1ZTX_LC 2B2X_LC 2CMR_LC	PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV	V LLNNFYPREAKVQWKVD V FLNNFYPRDISVKWKID V FLNNFYPRDISVKWKID V FLNNFYPRDINVKWKID V LLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID	NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI	DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLSLTKADYES DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER	IKVYACEVTHQGLS IKVYACEVTHQGLS INLYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST	SPVTKSFNR         211           SSPVTKSFNRGE         -         214           SSPVKSFNRNE         -         212           SSPVKSFNRNE         -         212           SSPVKSFNRNE         -         213           SSPVKS         -         208
2088_LC PTVSIFPPSSEQLTSGGASVVFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRNEC214 200N_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC214 20LPLC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC214 3085_LC PTVSIFPPSSEQLTSGGASVVFFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSLSSTLTLKBEYERHNSYTCEATHKTSTSPIVKSFNRGEC214 309A_LC PTVSIFPPSSEQLTSGGASVVFFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSLSSTLTLKBEYERHNSYTCEATHKTSTSPIVKSFNRGEC214 319J_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC214 116L1_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE214 118Z_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC214 118Z_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC214 118Z_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC214 118Z_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC214 118Y6_LC PTVSIFPPSSEQLTSGGASVVFLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC213 11Y9_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC213 11Y9_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC213 11Y9_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE213 11Y9_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE213 11Y9_LC PSVFIFPNSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE213 11Y9_LC PSVFIFPNSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTE	1LK3_LC 1MLC_LC 1OSP_LC 1UJ3_LC 1WEJ_LC 1YQV_LC 1ZTX_LC 2B2X_LC 2CMR_LC 2DD8_LC	PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVTIFPPSDEQLKSGTASV	V LLNNFYPREAKVQWKVD V FLNNFYPRDINVKWKID V LMNNFYPRDISVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V LLNNFYPREAKVQWKVD V LLNSFYPRAVTVAWKAD	NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI MALQSGNSQESVTEQI GSPVKAGVETTKPSKQ	DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLSLTKADYES DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK	IKVYACEVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST	SPVTKSFNR         211           SPVTKSFNRGE         214           SPVTKSFNRNE         214           SPVVKSFNRNE         214           'SPVVKSFNRNE         214           'SPVVKSFNRNE         214           'SPVKSFNRNE         211           'SPVKSFNRNE         212           'SPVKSFNRNE         212           'SPVKSFNRNE         213           'SPVTKSFNRNE         208           'SPVKTAPTECS         213
2QQN_LC PSVFIFPPSDEQLKSGTASVVLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE - 214 2R0L_LC PSVFIFPPSDEQLKSGTASVVLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE - 214 2VDR_LC PTVSIFPPSSEQLTSGGASVVCFLNNFYPRDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSLSSTLTLTKDEYERHNSYTCEATHKTSTSPTVKSFNRGE - 214 3D85_LC PSVFIFPPSDEQLKSGTASVVLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE - 214 1BJ1_LC PSVFIFPPSDEQLKSGTASVVLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE - 214 1CE1_LC PSVFIFPPSDEQLKSGTASVVLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE - 214 1N8Z_LC PSVFIFPPSDEQLKSGTASVVLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE - 214 1N8Z_LC PSVFIFPPSDEQLKSGTASVVLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE - 214 1SY6_LC PSVFIFPPSDEQLKSGTASVVLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE - 214 1N8Z_LC PSVFIFPPSDEQLKSGTASVVLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE - 214 1SY6_LC PSVFIFPPSDEQLKSGTASVVLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE - 214 1SY6_LC PSVFIFPPSDEQLKSGTASVVLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE - 214 1SY6_LC PSVFIFPPSDEQLKSGTASVVLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE - 214 1SY6_LC PSVFIFPPSDEQLKSGTASVVLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE - 213 1YY9_LC PSVFIFPPSDEQLKSGTASVVLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGA 213	1LK3_LC 1MLC_LC 1OSP_LC 1UJ3_LC 1WEJ_LC 1YQV_LC 2B2X_LC 2CMR_LC 2CMR_LC 2FD6_LC	PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQLTSGGASV	V LLNNFYPREAKVQWKVD V FLNNFYPRDINVKWKID V LNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V LLNNFYPRANVKKAD	NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSPVKAGVETTKPSKQ GSERQNGVLNSWTDQI	DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSMSSTLTLTKKDYER DSKDSTYSMSSTLSLTKADYES DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSLSSTLTLSKADYEK QS - NNKVAASSYLSLTPEQWKS DSKDSTYSMSSTLTLTKDEYER	IKVYACEVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST IKVYACEVTHQGLS IKSYSCQVTHEGST INSYTCEATHKTST	SPVTKSFNR         211           SPVTKSFNRBE         214           'SPIVKSFNRBE         211           'SPIVKSFNRBE         211           'SPIVKSFNRBE         213           'SPVKS         208           'SPVKSFNRBE         213
2R0L_LC       PSVFIFPPSDEQLKSGTASVVLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE       - 214         2VDR_LC       PTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTQDSKDSTYSLSSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRGE       - 214         3B85_LC       PSVFIFPPSDEQLKSGTASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE       - 214         3D9A_LC       PTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTQDSKDSTYSMSSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRGE       - 214         1BJ1_LC       PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE       - 214         1CE1_LC       PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE       - 211         1N8Z_LC       PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE       - 214         1SY6_LC       PTVSIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE       - 213         1N8Z_LC       PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE       - 214         1SY6_LC       PTVSIFPPSSEQLTSGGASVVCFLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE       - 213         1YY9_LC       PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGA       213	1LK3_LC 1MLC_LC 1OSP_LC 1UJ3_LC 1WEJ_LC 1YQV_LC 1ZTX_LC 2B2X_LC 2DD8_LC 2FD6_LC 2NXY_LC	PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PSVFIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVTLFPPSSEEFQANKATI PTVSIFPPSSEQLTSGGASV PSVFIFPPSDEQLKSGTASV	V LLNNFYPREAKVQWKVD V FLNNFYPRDISVKWKID V FLNNFYPRDISVKWKID V FLNNFYPRDINVKWKID V LLNNFYPREAKVQWKVD V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V LLNNFYPREAKVQWKVD V FLNNFYPREAKVQWKVD V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID	NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI NALQSGNSQESVTEQI	DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSMSSTLTLSKADYES DSKDSTYSMSSTLSLTKADYES DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER	IKVYACEVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST	SPVTKSFNR
2VDR_LC       PTVSIFPPSSEQLTSGGASVVFFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRBE214         3D85_LC       PSVFIFPPSDEQLKSGTASVVFLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRBE214         3D9A_LC       PTVSIFPPSSEQLTSGGASVVFLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSMSSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRBE214         1BJ1_LC       PSVFIFPPSDEQLKSGTASVVFLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRBE214         1CE1_LC       PSVFIFPPSDEQLKSGTASVVFLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE211         1N8Z_LC       PSVFIFPPSDEQLKSGTASVVFLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE214         1SY6_LC       PTVSIFPPSSEQLTSGGASVVFLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE213         1YY9_LC       PSVFIFPPSDEQLKSGTASVVFLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE213         1YY9_LC       PSVFIFPPSDEQLKSGTASVVFLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGA213	1LK3_LC 1MLC_LC 1OSP_LC 1UJ3_LC 1WEJ_LC 1YQV_LC 1ZTX_LC 2DA8_LC 2FD6_LC 2Q8B_LC	PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQLTSGGASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV	V LLNNFYPREAKVQWKVD V FLNNFYPRDISVKWKID V FLNNFYPRDISVKWKID V FLNNFYPRDINVKWKID V LLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD	NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI MALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI	DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLKDEYER DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLTKDEYER DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYER DSKDSTYSLSSTLTLSKADYER DSKDSTYSMSSTLTLTKDEYER	IKVYACEVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST	SPVTKSFNR         211           SSPVTKSFNRBE          214           SSPVKSFNRBE          214           SSPVKSFNRBE          214           SSPVKSFNRBE          214           SSPVKSFNRBE         -         214           SSPVKSFNRNE         -         211           SSPVKSFNRNE         -         212           SSPVKSFNRNE         -         213           SSPVKSFNRNE         -         213           SSPVKSFNRNE         213         -           SPVKSFNRNEA         214         -           SPVKSFNRNEA         214         -           SPVKSFNRNE         -         214
3D85_LC       PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE       - 214         3D9A_LC       PTVSIFPPSSEQLTSGGASVVCLLNNFYPRDINVKWKIDGSERQNOVLNSWTDQDSKDSTYSMSSTLTLTKDEYERHNSYTEATHKTSTSFIVKSFNRGE       - 213         1BJ1_LC       PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE       - 214         1CE1_LC       PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE       - 211         1N8Z_LC       PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE       - 214         1SY6_LC       PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE       - 213         1YY9_LC       PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGA       213         1YY9_LC       PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGA       213	1LK3_LC 1MLC_LC 1OSP_LC 1UJ3_LC 1WEJ_LC 1YQV_LC 2B2X_LC 2CMR_LC 2CD8_LC 2DD8_LC 2PD6_LC 2Q8B_LC 2Q0N_LC	PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV	V LLNNFYPREAKVQWKVD V FLNNFYPRDISVKWKID V FLNNFYPRDISVKWKID V FLNNFYPRDINVKWKID V LLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID	NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSPVKAGVETTKPSKQ GSERQNGVLNSWTDQI NALQSGNSQESVTEQI NALQSGNSQESVTEQI NALQSGNSQESVTEQI	DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLTKDEYER	IKVYACEVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST	SPVTKSFNR         211           SSPVTKSFNRGE         214           SSPVTKSFNRGE         214           SSPVKSFNRNE         214           SSPVKSFNRNE         214           SSPVKSFNRNE         214           SSPVKSFNRNE         214           SSPVKSFNRNE         214           SSPVKSFNRE         214           SSPVKSFNRE         214           SSPVKSFNRE         211           SSPVKSFNRE         212           SSPVKSFNRE         213           SSPVKSFNRE         213           SSPVKSFNRE         213           SSPVKSFNRE         213           SSPVKSFNRE         214
3D9ALC PTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERHNSYTCEATHRTSTSPIVKSFNRNE 213 1BJ1LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE 214 1CE1LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE 211 1N8Z_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE 214 1SY6_LC PTVSIFPPSSEQLTSGGASVVCFLNNFYPRDINVKWKIDGSERQNOVLNSWTDQDSKDSTYSLSSTLTLKKADYEKHKVYACEVTHQGLSSPVTKSFNRNEC 213 1YY9_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRNEC 213	1LK3_LC 1MLC_LC 1OSP_LC 1UJ3_LC 1WEJ_LC 1YQV_LC 1ZTX_LC 2CMR_LC 2CMR_LC 2DD8_LC 2FD6_LC 2Q8B_LC 2Q0N_LC 2Q0L_LC	PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV	V LLNNFYPREAKVQWKVD V FLNNFYPRDISVKWKID V FLNNFYPRDISVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V LISDFYPGAVTVAWKAD V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD	NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI NALQSGNSQESVTEQI NALQSGNSQESVTEQI NALQSGNSQESVTEQI NALQSGNSQESVTEQI NALQSGNSQESVTEQI	DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSMSSTLTLTKKDYER DSKDSTYSMSSTLSLTKADYES DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSLSSTLTLKDEYER DSKDSTYSLSSTLTLKDEYER DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYER DSKDSTYSLSSTLTLSKADYER DSKDSTYSLSSTLTLSKADYER	IKVYACEVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST IKVYACEVTHQGLS IKVYACEVTHQGLS	SPVTKSFNR
1BJ1_LC       PSVFIFPPSDEQLKSGTASVVLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE214         1CE1_LC       PSVFIFPPSDEQLKSGTASVVLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNR211         1N8Z_LC       PSVFIFPPSDEQLKSGTASVVLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE-214         1SY6_LC       PTVSIFPPSSEQLTSGGASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLKKADYEKHKVYACEVTHQGLSSPVTKSFNRNE213         1YY9_LC       PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLKADYEKHKVYACEVTHQGLSSPVTKSFNRGA213	1LK3_LC 1MLC_LC 1OSP_LC 1UJ3_LC 1WEJ_LC 1YQV_LC 1ZTX_LC 2DBX_LC 2DD8_LC 2FD6_LC 2QNLC 2R0L_LC 2VDR_LC	PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVLFPPSSEQLTSGGASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV	V LLNNFYPREAKVQWKVD V FLNNFYPRDISVKWKID V FLNNFYPRDISVKWKID V FLNNFYPRDINVKWKID V LLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V LLNNFYPRBAKVQWKVD V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRBAKVQWKVD V FLNNFYPRBAKVQWKVD V LLNNFYPRBAKVQWKVD V LLNNFYPRBAKVQWKVD V LLNNFYPRBAKVQWKVD	NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI MALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI	DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSMSSTLTLIKKDYER DSKDSTYSMSSTLSLTKADYES DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLTKDEYER DSKDSTYSLSSTLTLTKDEYER DSKDSTYSLSSTLTLTKDEYER DSKDSTYSLSSTLTLTKDEYER DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK	IKVYACEVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST IKVYACEVTHQGLS IKVYACEVTHQGLS IKVYACEVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTST	SPVTKSFNR.       211         SPVTKSFNRBE       214         SPIVKSFNRBE       214         SPVKSFNRBE       211         SPVKSFNRBE       212         SPVKSFNRBE       213         SPVKSFNRBE       213         SPVKSFNRBE       213         SPVKSFNRBE       213         SPVKSFNRBE       214         SPVKSFNRBE       213         SPVKSFNRBE       214         SPVKSFNRBE       213         SPVKSFNRBE       214
1CE1_LC       PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNR.       211         1N8Z_LC       PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGE       214         1SY6_LC       PTVSIFPPSSEQLTSGGASVVCFLNNFYPRDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERHNSYT EATHKTSTSFIVKSFNRHE       213         1YY9_LC       PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRHE       213         1YY9_LC       PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGA       213	1LK3_LC 1MLC_LC 1OSP_LC 1UJ3_LC 1WEJ_LC 1YQV_LC 1ZTX_LC 2DD8_LC 2DD8_LC 2FD6_LC 2QN_LC 2R0L_LC 2VDR_LC 3D85_LC	PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV	V LLNNFYPREAKVQWKVD V FLNNFYPRDISVKWKID V FLNNFYPRDISVKWKID V FLNNFYPRDINVKWKID V LLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPREAKVQWKVD V FLNNFYPREAKVQWKVD V FLNNFYPREAKVQWKVD V FLNNFYPREAKVQWKVD V FLNNFYPREAKVQWKVD V FLNNFYPREAKVQWKVD V FLNNFYPREAKVQWKVD	NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI MALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI NALQSGNSQESVTEQI NALQSGNSQESVTEQI NALQSGNSQESVTEQI NALQSGNSQESVTEQI NALQSGNSQESVTEQI	DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSMSSTLTLSKADYES DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLTKDEYER	IKVYACEVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST IKVYACEVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTST IKVYACEVTHQGLS	SPVTKSFNR       211         SSPVTKSFNRBE       214         SSPVTKSFNRBE       214         SSPVKSFNRBE       211         SSPVKSFNRNE       212         SSPVKSFNRBE       213         SSPVKSFNRBE       214
1N8Z_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAGEVTHQGLSSPVTKSFNRGE214 1SY6_LC PTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERHNSYTGEATHKTSTSPIVKSFNRNE213 1YY9_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYAGEVTHQGLSSPVTKSFNRGA213	1LK3_LC 1MLC_LC 1OSP_LC 1UJ3_LC 1WEJ_LC 1YQV_LC 1ZTX_LC 2DD8_LC 2DD8_LC 2FD6_LC 2QN_LC 2R0L_LC 2VDR_LC 3D85_LC	PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV	V LLNNFYPREAKVQWKVD V FLNNFYPRDISVKWKID V FLNNFYPRDISVKWKID V FLNNFYPRDINVKWKID V LLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V FLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD	NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI	DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK	IKVYACEVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST	SPVTKSFNR.       211         SPVTKSFNRGE       214         SPVKSFNRE       210         SPVKSFNRE       214         SPVKSFNRE       212         SPVKSFNRE       213         SPVKSFNRE       213         SPVKSFNRE       214
1SY6_LC PTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRNEC213 1YY9_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGA213	1LK3_LC 1MLC_LC 1OSP_LC 1UJ3_LC 1WEJ_LC 1YQV_LC 2B2X_LC 2CMR_LC 2DD8_LC 2CMR_LC 2Q8B_LC 2Q8B_LC 2Q0N_LC 2Q0N_LC 2VDR_LC 3D85_LC 3D9A_LC	PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV	V LLNNFYPREAKVQWKVD V FLNNFYPRDISVKWKID V FLNNFYPRDISVKWKID V FLNNFYPRDINVKWKID V LLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V FLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD	NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI	DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK	IKVYACEVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST	SPVTKSFNR.       211         SPVTKSFNRGE       214         SPVKSFNRE       210         SPVKSFNRE       214         SPVKSFNRE       212         SPVKSFNRE       213         SPVKSFNRE       213         SPVKSFNRE       214
1SY6_LC PTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDSTYSMSSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRNEC213 1YY9_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGA213	1LK3_LC 1MLC_LC 1OSP_LC 1UJ3_LC 1YQV_LC 1ZTX_LC 2DD8_LC 2CMR_LC 2QD8_LC 2Q8B_LC 2Q0N_LC 2Q0N_LC 2Q0N_LC 3D85_LC 3D9A_LC 1BJ1_LC	PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV	V LLNNFYPREAKVQWKVD V FLNNFYPRDISVKWKID V FLNNFYPRDISVKWKID V FLNNFYPRDINVKWKID V LLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V LLNNFYPREAKVQWKVD V FLNNFYPRDINVKWKID V LLNNFYPREAKVQWKVD V FLNNFYPREAKVQWKVD V FLNNFYPREAKVQWKVD V FLNNFYPREAKVQWKVD V FLNNFYPREAKVQWKVD V FLNNFYPREAKVQWKVD V FLNNFYPREAKVQWKVD V FLNNFYPREAKVQWKVD V FLNNFYPREAKVQWKVD	NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI NALQSGNSQESVTEQI NALQSGNSQESVTEQI NALQSGNSQESVTEQI NALQSGNSQESVTEQI NALQSGNSQESVTEQI NALQSGNSQESVTEQI NALQSGNSQESVTEQI NALQSGNSQESVTEQI NALQSGNSQESVTEQI NALQSGNSQESVTEQI	DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSMSSTLTLTKKDYER DSKDSTYSMSSTLSLTKADYES DSKDSTYSMSSTLSLTKADYES DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLKADYEK DSKDSTYSLSSTLTLTKDEYER DSKDSTYSLSSTLTLTKDEYER DSKDSTYSLSSTLTLTKDEYER DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER	IKVYACEVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST IKVYACEVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST	SPVTKSFNR
1YY9_LC PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGA 213	1LK3_LC 1MLC_LC 1SP_LC 1UJ3_LC 1WEJ_LC 1YQV_LC 1ZTX_LC 2DB8_LC 2CMR_LC 2CMR_LC 2QB8_LC 2QQN_LC 2QQN_LC 2QQN_LC 2QDR_LC 3D85_LC 3D85_LC 3D9A_LC 1CE1_LC	PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVLFPPSSEQLTSGGASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQLTSGGASV PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV	V LLNNFYPREAKVQWKVD V FLNNFYPRDISVKWKID V FLNNFYPRDISVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRBAKVQWKVD V FLNNFYPRBAKVQWKVD	NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI MALQSGNSQESVTEQI NALQSGNSQESVTEQI NALQSGNSQESVTEQI	DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLSLTKADYES DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYER DSKDSTYSLSSTLTLSKADYER DSKDSTYSLSSTLTLSKADYER DSKDSTYSLSSTLTLSKADYER DSKDSTYSLSSTLTLSKADYER DSKDSTYSLSSTLTLSKADYER DSKDSTYSLSSTLTLSKADYER DSKDSTYSLSSTLTLSKADYER DSKDSTYSLSSTLTLSKADYER DSKDSTYSLSSTLTLSKADYER DSKDSTYSLSSTLTLSKADYER	IKVYACEVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST IKVYACEVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST IKVYACEVTHQGLS	SPVTKSFNR.       211         SPVTKSFNRBE       214         SPIVKSFNRBE       211         SPIVKSFNRBE       213         SPVKSFNRBE       213         SPVKSFNRBE       213         SPVKSFNRBE       214         SPVKSFNRBE       213         SPVKSFNRBE       214         SPVKSFNRBE       214         SPVKSFNRE       214         SPVKSFNRBE       214         SPVKSFNRGE       214         SPVKSFNRGE       214         SPVKSFNRGE       214         SPVKSFNRGE       214         SPVKSFNRE       214
_ * * <b>* * * * * * * * *</b> * *	1LK3_LC 1MLC_LC 1SP_LC 1UJ3_LC 1WEJ_LC 1YQV_LC 1ZTX_LC 2DB8_LC 2DB8_LC 2FD6_LC 2R0N_LC 2R0N_LC 2R0N_LC 2R0N_LC 2R0N_LC 3D85_LC 3D85_LC 1BJ1_LC 1CE1_LC 1N8Z_LC	PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVSIFPPSSEQUTSGGASV PTVTIFPPSSEQUTSGGASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV	V LLNNFYPREAKVQWKVD V FLNNFYPRDISVKWKID V FLNNFYPRDISVKWKID V FLNNFYPRDINVKWKID V LLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V LLNNFYPREAKVQWKVD V FLNNFYPRDINVKWKID V LLNNFYPREAKVQWKVD V FLNNFYPREAKVQWKVD V FLNNFYPRDINVKWKID V LLNNFYPREAKVQWKVD V FLNNFYPRDINVKWKID V LLNNFYPREAKVQWKVD V FLNNFYPREAKVQWKVD V FLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD	NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI MALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI MALQSGNSQESVTEQI NALQSGNSQESVTEQI NALQSGNSQESVTEQI NALQSGNSQESVTEQI NALQSGNSQESVTEQI NALQSGNSQESVTEQI	DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSMSSTLTLSKADYES DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLTKDEYER DSKDSTYSLSSTLTLTKDEYER DSKDSTYSLSSTLTLTKDEYER DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK	IKVYACEVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST IKVYACEVTHQGLS IKVYACEVTHQGLS IKVYACEVTHQGLS IKVYACEVTHQGLS IKVYACEVTHQGLS	SPVTKSFNR.       211         SPVTKSFNR.       214         SPVTKSFNR.       214         SPVKSFNR.       210         SPVKSFNR.       214         SPVKSFNR.       215         SPVKSFNR.       211         SPVKSFNR.       211         SPVKSFNR.       213         SPVKSFNR.       214         SPVKSFNR.       213         SPVKSFNR.       214         SPVKSFNR.       213         SPVKSFNR.       214         SPVKSFNR.       214         SPVKSFNR.       213         SPVKSFNR.       213         SPVKSFNR.       214         SPVKSFNR.       213         SPVKSFNR.       213
	1LK3_LC 1MLC_LC 1OSP_LC 1UJ3_LC 1YQV_LC 2B2X_LC 2DD8_LC 2DD8_LC 2CMR_LC 2QN_LC 2Q8B_LC 2QQN_LC 2Q0N_LC 2VDR_LC 3D85_LC 3D9A_LC 1CE1_LC 1N8Z_LC 1SY6_LC	PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVSIFPPSSEQLTSGGASV PTVTIFPPSSEQLTSGGASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PTVSIFPPSSEQLTSGGASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV PSVFIFPPSDEQLKSGTASV	V LLNNFYPREAKVQWKVD V FLNNFYPRDISVKWKID V FLNNFYPRDISVKWKID V FLNNFYPRDINVKWKID V LLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPRDINVKWKID V FLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V FLNNFYPRDINVKWKID V FLNNFYPREAKVQWKVD V FLNNFYPREAKVQWKVD V FLNNFYPREAKVQWKVD V FLNNFYPREAKVQWKVD V FLNNFYPREAKVQWKVD V FLNNFYPREAKVQWKVD V FLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD V LLNNFYPREAKVQWKVD	NALQSGNSQESVTEQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI MALQSGNSQESVTEQI GSERQNGVLNSWTDQI NALQSGNSQESVTEQI GSERQNGVLNSWTDQI	DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSMSSTLTLSKADYEK DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSMSSTLTLTKDEYER DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK DSKDSTYSLSSTLTLSKADYEK	IKVYACEVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST IKVYACEVTHQGLS INSYTCEATHKTST IKVYACEVTHQGLS IKVYACEVTHQGLS IKVYACEVTHQGLS IKVYACEVTHQGLS IKVYACEVTHQGLS IKVYACEVTHQGLS IKVYACEVTHQGLS INSYTCEATHKTST	SPVTKSFNR.         211           SSPVTKSFNR.         214           SSPVTKSFNR.         214           SSPVKSFNR.         210           SSPVKSFNR.         210           SSPVKSFNR.         214           SSPVKSFNR.         214           SSPVKSFNR.         214           SSPVKSFNR.         214           SSPVKSFNR.         214           SSPVKSFNR.         214           SSPVKSFNR.         212           SSPVKSFNR.         212           SSPVKSFNR.         212           SSPVKSFNR.         213           SSPVKSFNR.         214           SSPVKSFNR.         214

Fig. 2 Multiple sequence alignments for (a) light and (b) heavy chains of the Fabs in our dataset. The Fab sequences are arranged in the same order as they are listed in Table I. All conserved Cys residues are highlighted in green. The CDRs are highlighted in yellow. The predicted APRs are shown in red letters. Kabat numbering is shown at the bottom of sequence alignment (39).

of interface is  $56 \pm 7\%$  polar. This value is similar to an earlier average value (49%) for antibody-antigen interfaces (35). The electrostatic and polar interactions have been known to be important factors in determining affinity and specificity of antibody-antigen complexes (47,48).

The APRs considered in the Fig. 4b are those found in the antigen-binding regions of Fabs. The average polar fraction of buried surface area of APRs is  $55\pm19\%$  (range 9–93%) after excluding two outliers described below. This value is

🖄 Springer

similar to that for Fabs. However, the variation is now wider. In case of murine antibody Fab-Protein A complex (Complex 9), the polar fraction for this complex is not available because none of the APRs contribute to the buried surface areas. For the Fab NMC4-Von Willebrand factor complex (Complex 2), the buried surface area of the APRs comes from only one binding residue (E92), which is hydrophilic, leading to 100% polar fraction. In four out of the five commercial Fab-antigen complexes, the APRs show greater than 50% polar fraction.

b

1FE8 HC	DVKLVQSGPGLVAPSQSLSIT	CTVS	3FSLTTYGVS	VROPPGKGLEWLG	VIWGDGNTTYHSA	TSRLSTSKDNSRSOVFLK	LNSLHTDDTATYY	AGNYYGM	-DYW 105	5
1FNS HC	QVQLKESGPGLVAPSQSLSIT									
1HOD HC	EVMLVESGGGLVKPGGSLKLS									
1IQD HC	QVQLVQSGAEVKKPGASVKVS									
1JPS HC	EVQLVESGGGLVQPGGSLRLS									
1KB5 HC	EVQLQQSGPELEKPGASVKIS									
1LK3 HC	QVNLLQSGAALVKPGASVKLS									
1MLC HC	QVQLQESGAEVMKPGASVKIS									
10SP HC	EVQLQESGPSLVKPSQTLSLT									
1UJ3 HC	QVQLLESGAVLARPGTSVKIS									
1WEJ HC	EVQLQQSGAELVKPGASVKLS	CTAS	GFNIKDTYMH	WVKQRPEKGLEWIG	RIDPASGNTKYDPK	QDKATITADTSSNTAYLQ	lssltsedtavyy <mark>c</mark>	AG <mark>YDYGN</mark>	<mark>IFDY</mark> W 107	7
1YQV HC	EVQLQQSGAELMKPGASVKIS	CKAS	GYTFSDYWIE	WVKQRPGHGLEWIG	EILPGSGSTNYHER	KGKATFTADTSSSTAYMQ	LNSLTSEDSG <mark>VYY</mark> C	LH <mark>GNYDF</mark>	-DGW 106	5
1ZTX HC	QVQLQQSGSELMKPGASVQIS	CKAT	GYTFSDYWIE	WVKQRPGHGLEW <mark>IG</mark>	DILCGTGRTRYNEK		LSSLTSEDS <mark>AVYY</mark> C	AR <mark>SASYGDYA</mark>	-DYW 109	9
2B2X HC	EVQLVESGGGLVQPGGSLRLS	CAAS	GFTFSRYTMS	WVRQAPGKGLE <mark>WVA</mark>	VISGGGHTYYLDS	<mark>/EG</mark> RFTISRDNSKNTLYLQI	MNSLRAEDT <mark>AVYY</mark>	TR <mark>GFGDGG</mark> Y	FDV <mark>W 108</mark>	3
2 CMR HC	QVQLVQSGAEVRKPGASVKVS	CKAS	GDTFSSYAIS	VRQAPGQGLEWMG	GIIPIFGTANYAQA	QGRVTITANESTSTAYME	LSSLRSEDT <mark>AIYY</mark>	AR <mark>DNPTLLGS</mark>	-DYW 109	Э
2DD8 HC	QVQLQQSGAEVKKPGSSVKVS	CKAS	GGTFSSYTIS	<b>VV</b> RQAPGQGLEWMG	GITPILGIANYAQK	QGRVTITTDESTSTAYME	LSSLRSEDT <mark>AVYY</mark>	AR <mark>DTVMGGM</mark>	<mark>-DV</mark> W 108	3
2FD6_HC	GVKLQQSGPEVVKPGASVKIS									
2NXY_HC	EVQLVESGAEVKKPGSSVKVS									3
2Q8B_HC	EVQLQQSGAELLRPGASVKLS									
2QQN_HC	EVQLVESGGGLVQPGGSLRLS									
2R0L_HC	EVQLVESGGGLVQPGGSLRLS									
2VDR_HC	EVQLQQSGAELVKPGASVKLS									
3D85_HC	EVQLQQSGPELVKPGASVKMS									5
3D9A_HC	DVQLQESGPSLVKPSQTLSLT									3
1BJ1_HC	EVQLVESGGGLVQPGGSLRLS									
1CE1_HC	QVQLQESGPGLVRPSQTLSLT									L
1N8Z_HC	EVQLVESGGGLVQPGGSLRLS									)
1SY6_HC	QVQLQQSGAELARPGASVKMS									
1YY9_HC	QVQLKQSGPGLVQPSQSLSIT	CTVS	GFSLTNYGVH	WVRQSPGKGLEWLG	VIWSGGNTDYNTP * ****		MNSLQSNDTAIYY	AR <mark>ALTYYDYB</mark> *********	FAYW 109	Э
	10 20		30	40	50 ABC3 60		ABC 90	OABCDEFGHIJ		
1FE8_HC	GQGTSVTVSSAETTAPSVYKL									)
1FNS_HC	GQGTSVTVSSAKTTPPSVYPL	APGS	AAQTNSMVT <mark>L</mark>	G <mark>C</mark> LVKGYFPEPVTV	TWNSGSLSSGVHTFPA	LQS-DLYTLSSSVTVPSS	twpsetvt <mark>c</mark> nvahe	ASSTKVDKKIVPRDCG	225	5
1FNS_HC 1H0D_HC	GQGTSVTVSSAKTTPPSVYPL GQGTSVTVSSAKTTPPSVYPL	APGS APGG	AAQTNSMVT <mark>L</mark> ( 3GGGGAM <mark>VTL</mark> (	G <mark>CLV</mark> KGYFPEPVTV G <mark>C</mark> LVKGYFPEPVTV	TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA	/LQS-DLYTLSSSVTVPSS /LQS-DLYTLSSSVTVPSS	TWPSETVT <mark>C</mark> NVAHE PWPSETVT <mark>C</mark> NVAHE	PASSTKVDKKIVPRDCG PASSTKVDKKIVPRDC	225	5
1FNS_HC 1H0D_HC 1IQD_HC	GQGTSVTVSSAKTTPPSVYPL GQGTSVTVSSAKTTPPSVYPL GQGTMVTVSSASTKGPSVFPL	APGS APGG APCS	AAQTNSMVTL( GGGGGAM <mark>VTL(</mark> RSTSESTAAL(	3 <mark>C</mark> LVKGYFPEPVTV 3CLVKGYFPEPVTV 3 <mark>C</mark> LVKDYFPEPVTV	TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA SWNSGALTSGVHTFPA	/LQS-DLYTLSSSVTVPSS /LQS-DLYTLSSSVTVPSS /LQSSGLYSLSSVVTVPSS	TWPSETVT <mark>C</mark> NVAHE PWPSETVTCNVAHE SLGTATYT <mark>C</mark> NVDHE	PASSTKVDKKIVPRDCG PASSTKVDKKIVPRDC VPSNTKVDKRV	225 223 215	5
1FNS_HC 1HOD_HC 1IQD_HC 1JPS_HC	GQGTSVTVSSAKTTPPSVYPL GQGTSVTVSSAKTTPPSVYPL GQGTMVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL	APGS APGG APCS APSS	AAQTNSMVTL( GGGGGAMVTL( RSTSESTAAL( KSTSGGTAAL(	G <mark>CLVKGYFPEPVTV GCLVKGYFPEPVTV GCLVKDYFPEPVTV GCLVKDYFPEPVTV</mark>	'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA	/LQS-DLYTLSSSVTVPSS /LQS-DLYTLSSSVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS	TWPSETVT <mark>C</mark> NVAHE PWPSETVTCNVAHE SLGTATYTCNVDHK SLGTQTYICNVNHK	PASSTKVDKKIVPRDCG PASSTKVDKKIVPRDC KPSNTKVDKRV KPSNTKVDKKVEPKSCDKT	225 223 215 THT 225	5
1FNS_HC 1H0D_HC 1IQD_HC 1JPS_HC 1KB5_HC	GQGTSVTVSSAKTTPPSVYPL GQGTSVTVSSAKTTPPSVYPL GQGTMVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLLVSSAKTTAPSVYPL	APGS APGG APCS APSS APSS	AAQTNSMVTLO GGGGGAMVTLO RSTSESTAALO KSTSGGTAALO GDTTGSSVTLO	GCLVKGYFPEPVTV GCLVKGYFPEPVTV GCLVKDYFPEPVTV GCLVKDYFPEPVTV GCLVKGYFPEPVTL	TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA SWNSGALTSGVHTFPA SWNSGALTSGVHTFPA TWNSGSLSSGVHTFPA	/LQS-DLYTLSSSVTVPSS /LQS-DLYTLSSSVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS /LQS-DLYTLSSSVTVTSS	TWPSETVT <mark>C</mark> NVAHE PWPSETVTCNVAHE SLGTATYTCNVDHK SLGTQTYICNVNHK TWPSQSIT <mark>C</mark> NVAHE	PASSTKVDKKIVPRDCG PASSTKVDKKIVPRDC KPSNTKVDKRV KPSNTKVDKKVEPKSCDKT PASSTKVDKKIEPR	225 223 215 THT 225 219	5
1FNS_HC 1HOD_HC 1IQD_HC 1JPS_HC 1KB5_HC 1LK3_HC	GQGTSVTVSSAKTTPPSVYPL GQGTSVTVSSAKTTPPSVYPL GQGTMVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLTVSSAKTTAPSVYPL GQGTLVTVSSAETTAPSVYPL	APGS APGG APCS APSS APSS APVC	AAQTNSMVTLO GGGGGAMVTLO RSTSESTAALO KSTSGGTAALO GDTTGSSVTLO ALKSNSMVTLO	CLVKGYFPEPVTV LVKGYFPEPVTV LVKDYFPEPVTV LVKDYFPEPVTV LVKGYFPEPVTL LVKGYFPEPVTL	TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA SWNSGALTSGVHTFPA SWNSGALTSGVHTFPA TWNSGSLSSGVHTFPA TWNSGALSSGVHTFPA	/LQS-DLYTLSSSVTVPSS' /LQS-DLYTLSSSVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS /LQS-DLYTLSSSVTVTSS' /LQS-GLYTLTSSVTVPSS'	TWPSETVT <mark>C</mark> NVAHE PWPSETVTCNVAHE SLGTATYTCNVDHK SLGTQTYICNVNHK TWPSQSITCNVAHE TWPSQTVTCNVAHE	PASSTKVDKKIVPRDCG PASSTKVDKKIVPRDC (PSNTKVDKRV (PSNTKVDKKVEPKSCDKT PASSTKVDKKIEPR PASSTKVDKKIVPR	225 223 215 HT 225 219 219	5
1FNS_HC 1HOD_HC 1IQD_HC 1JPS_HC 1KB5_HC 1LK3_HC 1MLC_HC	GQGTSVTVSSAKTTPPSVYPL GQGTSVTVSSASKTPPSVYPL GQGTWVVSSASTKGPSVFPL GQGTLVVSSASTKGPSVFPL GQGTLVVSSASTTAPSVYPL GQGTLVVSSASTTAPSVYPL	APGS APGG APCS APSS APVC APGT APGS	AAQTNSMVTL( GGGGGAM <mark>VTL</mark> ( RSTSESTAAL( KSTSGGTAAL( GDTTGSSVTL( ALKSNSM <mark>VTL</mark> ( AAQTNSM <mark>VTL</mark> (	UKGYFPEPUTV LVKGYFPEPUTV LVKDYFPEPUTV LVKDYFPEPUTV LVKGYFPEPUTL JVKGYFPEPUTV LVKGYFPEPUTV	"TWNSGSLSSGVHTFPA "TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA TWNSGSLSSGVHTFPA 'TWNSGALSSGVHTFPA 'TWNSGSLSSGVHTFPA	/LQS-DLYTLSSSVTVPSS' /LQS-DLYTLSSSVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS /LQS-DLYTLSSSVTVTSS' /LQS-GLYTLTSSVTVPSS' /LQS-DLYTLSSSVTVPSS	TWPSETVT <mark>C</mark> NVAHE PWPSETVTCNVAHE SLGTATYTCNVDHK SLGTQTYICNVHK TWPSQSITCNVAHE TWPSQTVTCNVAHE PRPSETVTCNVAHE	ASSTKVDKKIVPRDCG PASSTKVDKKIVPRDC CPSNTKVDKKVEPKSCDKI PASSTKVDKKIEPR PASSTKVDKKIVPR PASSTKVDKKIVPRDC	225 223 215 HT 225 219 219 218	5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
1FNS_HC 1H0D_HC 1IQD_HC 1JPS_HC 1KB5_HC 1LK3_HC 1MLC_HC 1OSP_HC	GQGTSVTVSSAKTTPPSVYPL GQGTSVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTTAPSVYPL GQGTLVTVSSASTTPPSVPPL GQGTLTVTVSASTTPPSVPPL	APGS APGG APCS APSS APVC APGT APGS	AAQTNSMVTL GGGGGAMVTL RSTSESTAAL KSTSGGTAAL GDTTGSSVTL ALKSNSMVTL AAQTNSMVTL GDTTGSSVTL	LVKGYFPEPVTV CLVKGYFPEPVTV LVKDYFPEPVTV LVKDYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV	"TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSSVHTFPA	/LQS-DLYTLSSSVTVPSS' /LQS-DLYTLSSSVTVPSS: /LQSSGLYSLSSVVTVPSS: /LQS-DLYTLSSSVTVPSS: /LQS-CLYTLTSSVTVPSS' /LQS-DLYTLSSSVTVPSS' /LQS-DLYTLSSSVTVPSS'	TWPSETVT <mark>O</mark> NVAHF PWPSETVTONVAHF SLGTATYTONVHK SLGTQTYIONVHK TWPSQSITONVAHF TWPSQTVTONVAHF PRPSETVTONVAHF TWPSQTVTOSVAHF	PASSTKVDKKIVPRDC PASSTKVDKKIVPRDC (PSNTKVDKRV PASSTKVDKKVEPKSCDKT PASSTKVDKKIEPR PASSTKVDKKIVPRC PASSTVVDKKLE	225 223 215 215 219 219 218 218	5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
1FNS_HC 1H0D_HC 1JQD_HC 1JPS_HC 1KB5_HC 1LK3_HC 1MLC_HC 10SP_HC 1UJ3_HC	GQGTSVTVSSAKTTPPSVYPL GQGTSVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTTAPSVYPL GQGTLVTVSSASTTPSVVFL GQGTLVTVSSASTTPPSVVFL GQGTLVTVSSASTKGPSVFPL	APGS APGG APCS APSS APVC APGT APGS APGS APGC	AAQTNSMVTL GGGGGAMVTL RSTSESTAAL KSTSGGTAAL GDTTGSSVTL ALKSNSMVTL AAQTNSMVTL GDTTGSSVTL RSTSESTAAL	LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPESVTV LVKGYFPESVTV	"TWNSGSLSSGVHTFPA "TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA JSWNSGALTSGVHTFPA "TWNSGSLSSGVHTFPA "TWNSGSLSSGVHTFPA "TWNSGSLSSSVHTFPA 'SWNSGALTSGVHTFPA	/LQS-DLYTLSSSVTVPSS' /LQSSGLYSLSSVVTVPSS: /LQSSGLYSLSSVVTVPSS: /LQS-DLYTLSSSVTVPSS' /LQS-GLYTLTSSVTVPSS' /LQS-GLYTLTSSVTVPSS' /LQS-GLYTMSSSVTVPSS' /LQSSGLYSLSSVVTVPSS'	TWPSETVT <mark>O</mark> NVAHE PWPSETVTONVAHE SLGTATYTONVDHK SLGTQTYIONVAHE TWPSQSITONVAHE PRPSETVTONVAHE PRPSETVTONVAHE SLGTKTYT <mark>O</mark> NVDHK	ASSTKVDKKIVPRDCG ASSTKVDKKIVPRDC (PSNTKVDKRV PSNTKVDKKVEPRSCDKT ASSTKVDKKIEPR ASSTKVDKKIVPR ASSTKVDKKIPR (PSNTKVDKRLE	225 223 215 -HT 225 219 218 218 218	5 5 5 5 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9
1FNS_HC 1H0D_HC 1JQD_HC 1JPS_HC 1KB5_HC 1KK3_HC 1MLC_HC 10SP_HC 1UJ3_HC 1WEJ_HC	GQGTSVTVSSAKTTPPSVYPL GQGTSVTVSSAKTPPSVYPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVVSSASTTPSVYPL GQGTLVVSSASTTPPSVYPL GQGTLVVSSASTTPPSVYPL GQGTLVVSSASTKGPSVFPL GQGTLTVSSASTKGPSVFPL	APGS, APGG APCS APSC APVC APGT APGS APGC APCS APCS	AAQTNSMVTL GGGGGAMVTL RSTSESTAAL KSTSGGTAAL GDTTGSSVTL ALKSNSMVTL GDTTGSSVTL RSTSESTAAL AALKSSMVTL	LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPESVTV LVKGYFPESVTV LVKGYFPEPVTV	TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA	/LQS-DLYTLSSSVTVPSS /LQS-DLYTLSSSVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS /LQS-DLYTLSSSVTVPSS /LQS-DLYTLSSVTVPSS /LQS-GLYTMSSSVTVPSS /LQSSGLYSLSSVVTVPSS /LQS-DLYTLTSSVTVPSS	TWPSETVT <mark>C</mark> NVAHE PWPSETVTCNVAHE SLGTATYTCNVDHK SLGTQTYLCNVNHK TWPSQSITCNVAHE TWPSQTVTCNVAHE TWPSQTVTCSVAHE SLGTKTYCCNVAHE TWPSQTVTCNVAHE	PASSTKVDKKIVPRDCG PASSTKVDKKIVPRDC CPSNTKVDKKV= PASSTKVDKKIEPR PASSTKVDKKIVPR PASSTKVDKKIVPRDC PASSTKVDKKLE CPSNTKVDKKLE	225 223 215 219 219 218 218 217 OC 223	5 5 5 5 9 9 9 3 3 7 3
1FNS_HC 1H0D_HC 1IQD_HC 1JPS_HC 1KB5_HC 1KK3_HC 1MLC_HC 10SP_HC 1UJ3_HC 1WEJ_HC 1YQV_HC	GQGTSVTVSSAKTTPPSVYPL GQGTSVTVSSASTKTPPSVYPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTTAPSVYPL GQGTLVTVSSASTTPPSVYPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLTVSSASTTPPSVYPL	APGS APGG APCS APSS APVC APGT APGS APGS APGT APGS	AAQTNSMVTLG GGGGGAMVTL RSTSESTAALG SDTTGSSVTL ALKSNSMVTLG AAQTNSMVTLG 3DTTGSSVTL RSTSESTAALG AALKSSMVTL AAQTNSMVTLG	LVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV	TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA	/LQS-DLYTLSSSVTVPSS' /LQS-DLYTLSSSVTVPSS: /LQSSGLYSLSSVVTVPSS: /LQS-GLYTLSSSVTVPSS' /LQS-GLYTLSSSVTVPSS' /LQS-GLYTLSSSVTVPSS' /LQS-GLYTMSSSVTVPSS' /LQS-DLYTLSSVTVPSS' /LQS-DLYTLSSVTVPSS'	TWPSETVTONVAHE PWPSETVTONVAHE SLGJTQTYINVDH SLGJQTYIONVNH TWPSQSITONVAHE PRPSETVTONVAHE TWPSQTVTOSVAHE SLGTKTYTONVAHE TWPSQTVTONVAHE	ASSTKVDKKIVPRDC ASSTKVDKKIVPRDC CPSNTKVDKKVEPKSCDKT ASSTKVDKKIEPR ASSTKVDKKIVPRC ASSTKVDKKIVPRDC ASSTKVDKKIVPRDC ASSTKVDKKIVPRNCGGE ASSTKVDKKIVDR	225 215 :HT 225 219 219 218 218 217 OC 223 215	- 5 3 5 5 9 9 9 3 3 7 3 5
1FNSHC 1H0DHC 1JQDHC 1JPSHC 1KB5HC 1KK3HC 1MLCHC 10SPHC 1UJ3HC 1WEJHC 1YQVHC 1ZTXHC	GQGTSVTVSSAKTTPPSVYPL GQGTSVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTTAPSVYPL GQGTLVTVSSASTTPPSVYPL GQGTLVTVSSASTTPPSVYPL GQGTLVTVSSASTTPPSVYPL GQGTLLTVSSASTTPPSVYPL GQGTLLTVSSASTTPPSVYPL	APGS APGS APCS APSS APVC APGT APGS APGS APGT APGS APGS	AAQTNSMVTLG GGGGGAMVTLG KSTSGGTAALG KSTSGGTAALG AAQTNSMVTLG JDTTGSSVTLG AAQTNSMVTLG AALKSSMVTLG AALKSSMVTLG AAQTNSMVTLG GDTTGSSVTLG	LVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTU SLVKGYFPEPVTU SLVKGYFPEPVTU SLVKGYFPEPVTV SLVKGYFPESVTV SLVKGYFPESVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPESVTV	TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA 'TWNSGLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA	/LQS-DLYTLSSSVTVPSS' /LQSSGLYSLSSVVTVPSS: /LQSSGLYSLSSVVTVPSS: /LQS-DLYTLSSSVTVPSS: /LQS-CLYTLSSSVTVPSS' /LQS-CLYTLSSSVTVPSS' /LQS-CLYTMSSSVTVPSS' /LQSSGLYSLSSVVTVPSS' /LQS-DLYTLTSSVTVPSS' /LQS-DLYTLSSSVTVPSS' /LQS-GLYTMSSSVTVPSS'	TWPSETVTONVAHE PWPSETVTONVAHE SLGTATYTONVDHK SLGTQTYIONVNHK TWPSQSITONVAHE TWPSQTVTONVAHE TWPSQTVTONVAHE SLGTKTYTONVAHE TWPSQTVTONVAHE TWPSQTVTONVAHE TWPSQTVTONVAHE	ASSTKVDKKIVPRDC ASSTKVDKKIVPRDC (PSNTKVDKRV ASSTKVDKKIEPR ASSTKVDKKIVPR ASSTKVDKKIVPRDC (PSNTKVDKKLE PASSTKVDKKIVPRNCGGL ASSTTVDKKLEPS PASSTKVDKKILP	225 215 :HT 225 219 219 218 218 217 20C 223 215 219	5 5 5 5 9 9 9 3 5 7 3 5 9 7 3 5 9
1FNS_HC 1H0D_HC 1JPS_HC 1KB5_HC 1KK3_HC 1MLC_HC 1OSP_HC 1UJ3_HC 1WEJ_HC 1ZTX_HC 2B2X_HC	GQGTSVTVSSAKTTPPSVYPL GQGTSVTVSSAKTTPPSVYPL GQGTLVTVSSASTKGPSVFPL GQGTLIVSSASTKGPSVFPL GQGTLIVSSASTTPSVYPL GQGTLIVSSASTTPPSVYPL GQGTLIVSSASTTPPSVYPL GQGTLIVSSASTTPPSVYPL GQGTLIVSSASTTPPSVYPL GQGTLIVSSAKTTPPSVYPL GQGTLIVSSAKTTPPSVYPL GQGTLIVSSAKTPPSVYPL	APGS APGG APCS APSS APS APGT APGS APGS APGS APGS APGS	AAQTNSMVTLG GGGGGMVTL KSTSGTAAL GDTTGSSVTLG AAQTNSMVTL AAQTNSMVTL AAQTNSMVTLG AALKSSMVTL AAQTNSMVTLG AAQTNSMVTLG	LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEVTV	TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA	/LQS-DLYTLSSSVTVPSS: /LQS-DLYTLSSSVTVPSS: /LQSSGLYSLSSVVTVPSS: /LQSSGLYSLSSVVTVPSS: /LQS-DLYTLSSSVTVPSS: /LQS-CLYTMSSSVTVPSS: /LQS-CLYTMSSSVTVPSS: /LQS-DLYTLSSVTVPSS: /LQS-DLYTLSSVTVPSS: /LQS-DLYTLSSVTVPSS: /LQS-DLYTLSSSVTVPSS:	TWPSETVTONVAHE PWPSETVTONVAHE SLGTATYT NVDHK SLGTQTY I ONVNHK TWPSQSITONVAHE PRPSETVTONVAHE TWPSQTVTONVAHE TWPSQTVTONVAHE TWPSQTVTONVAHE TWPSQTVTONVAHE TWPSQTVTONVAHE TWPSQTVTONVAHE TWPSQTVTONVAHE TWPSQTVTONVAHE	PASSTKVDKKIVPRDCG PASSTKVDKKIVPRDC VPSNTKVDKKVEPKSCDKT VPSNTKVDKKIPR PASSTKVDKKIVPR PASSTKVDKKIVPRDC PASSTKVDKKIVPROCGGI PASSTKVDKKIVPROCGGI PASSTKVDKKIVPROCGHH	225 223 215 'HT 225 219 218 218 218 217 OC 223 219 'HHH- 226	5 5 5 5 7 3 5 7 3 5 7 3 5 7 3 5 7 5 7 5
1FNS_HC 1HOD_HC 1JQD_HC 1JPS_HC 1KB5_HC 1LK3_HC 1MLC_HC 1OSP_HC 1UJ3_HC 1WEJ_HC 1YQV_HC 1ZTX_HC 2B2X_HC 2CMR_HC	GQGTSVTVSSAKTTPPSVYPL GQGTSVTVSSASTKPSVFPL GQGTWVVSSASTKGPSVFPL GQGTLVVSSASTKGPSVFPL GQGTLVVSSASTTAPSVYPL GQGTLIVSSASTTPPSVPPL GQGTLVVSSASTKGPSVFPL GQGTLVVSSASTKGPSVFPL GQGTLIVSSASTTPPSVYPL GQGTLIVSSASTTPPSVYPL GQGTLVVSSASTTPPSVYPL GQGTLVVSSASTTPPSVYPL GQGTLVVSSASTTPPSVYPL	APGS APGG APCS APS APGT APGT APGC APGC APGT APGS APGS APGS APGS	AAQTNSMVTLG GGGGGAWVTLG KSTSGTAALG SDTGSSVTLG AAQTNSMVTLG AAQTNSMVVTLG SDTGSSVTLG AALKSSMVTLG AALKSSMVTLG AAQTNSMVTLG SDTGSSVTLG AAQTNSMVTLG	CLVKGYFPEPVTV CLVKGYFPEPVTV CLVKGYFPEPVTV LVKGYFPEPVTV CLVKGYFPEPVTV CLVKGYFPEPVTV CLVKGYFPEPVTV CLVKGYFPEPVTV CLVKGYFPEPVTV CLVKGYFPESVTV CLVKGYFPESVTV CLVKGYFPEPVTV CLVKGYFPEPVTV	TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA	/LQS-DLYTLSSSVTVPSS' /LQS-DLYTLSSSVTVPSS' /LQSSGLYSLSSVVTVPSS' /LQSSGLYSLSSVVTVPSS' /LQS-DLYTLSSSVTVPSS' /LQS-OLYTLSSSVTVPSS' /LQS-GLYTMSSSVTVPSS' /LQS-DLYTLSSVTVPSS' /LQS-CLYTMSSSVTVPSS' /LQS-CLYTMSSSVTVPSS' /LQS-DLYTLSSSVTVPSS'	TWPSETVTONVAHE PWPSETVTONVAHE SLGTQTYIONVAHE TWPSQSITONVAHE TWPSQTVTONVAHE TWPSQTVTONVAHE TWPSQTVTONVAHE TWPSCTVTONVAHE TWPSETVTONVAHE TWPSETVTONVAHE SLGTQTYIONVAHE SLGTQTYIONVAHE	PASSTKVDKKIVPRDC PASSTKVDKKIVPRDC PASSTKVDKKVEPKSCDKT PASSTKVDKKIVPR PASSTKVDKKIVPRC PASSTKVDKKIVPRDC PASSTKVDKKIVPRNCGGE PASSTKVDKKILD PASSTKVDKKILD PASSTKVDKKIVPRCHHF PASSTKVDKRV	225 215 HT 219 219 218 218 217 0C 223 215 215 215 217	- 5 3 5 5 9 9 9 3 3 7 3 5 9 5 7 5 9 7
1FNS <sup>-</sup> HC 1HQD <sub>-</sub> HC 1JQP <sub>3</sub> HC 1KB5 <sup>-</sup> HC 1KK3 <sub>-</sub> HC 1MLC <sub>-</sub> HC 10SP <sub>-</sub> HC 1UJ3 <sub>-</sub> HC 1WQJ <sub>-</sub> HC 1YQV <sub>-</sub> HC 2ZX <sub>-</sub> HC 2CMR <sub>-</sub> HC 2DD8 <sub>-</sub> HC	GQGTSVTVSSAKTTPPSVYPL GQGTSVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTTAPSVYPL GQGTLVTVSSASTTAPSVYPL GQGTLVTVSSASTKGPSVFPL GQGTLTVSSASTKGPSVFPL GQGTLTVSSAKTTPPSVYPL GQGTLTVSSAKTTPPSVYPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL	APGS APGG APCS APVC APGT APGS APGC APGS APGS APGS APGS APGS APGS	AAQTNSMVTLG GGGGGAWTLL KSTSGTAALG KSTSGTAALG STGGSSVTL ALKSNSMVTL GDTTGSSVTLG AAQTNSMVTLG GDTTGSSVTLG AAQTNSMVTLG GDTTGSSVTLG AAQTNSMVTLG KSTSGGTAALG	LVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV	TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA	/LQS-DLYTLSSSVTVPSS' /LQS-DLYTLSSSVTVPSS: /LQSSGLYSLSSVVTVPSS: /LQS-GLYTLSSSVTVPSS' /LQS-GLYTLSSSVTVPSS' /LQS-GLYTLSSSVTVPSS' /LQS-GLYTMSSSVTVPSS' /LQS-DLYTLSSVTVPSS' /LQS-GLYTMSSSVTVPSS' /LQS-GLYTMSSSVTVPSS' /LQS-GLYTLSSSVTVPSS' /LQSSGLYSLSSVVTVPSS' /LQSSGLYSLSSVVTVPSS'	TWPSETVTONVAHE PWPSETVTONVAHE SLGJATYTINVDHK SLGJATYTINVDHK TWPSQSITONVAHE TWPSQTVTONVAHE TWPSQTVTONVAHE TWPSQTVTONVAHE TWPSETVTONVAHE TWPSETVTONVAHE TWPSETVTONVAHE SLGJATYTINVAHE SLGJATYTINVAHE	ASSTKVDKKIVPRDC ASSTKVDKKIVPRDC (PSNTKVDKKVEPKSCDKT PASSTKVDKKIVPR PASSTKVDKKIVPRC PASSTKVDKKIVPRDC PASSTKVDKKIVPRCGGI PASSTKVDKKILD PASSTVDKKLEPS PASSTKVDKKILPS	225 223 219 219 218 218 217 217 215 219 HHHH - 226 SPLF 228	5 5 5 5 5 5 9 9 3 5 7 5 7 5 7 3
1FNS_HC 1HOD_HC 1IQD_HC 1JPS_HC 1KB5_HC 1LK3_HC 1MLC_HC 1OSP_HC 1UJ3_HC 1UJ3_HC 1VEJ_HC 1ZTX_HC 2EX_HC 2CMR_HC 2FD6_HC	GQGTSVTVSSAKTTPPSVYPL GQGTSVTVSSAKTPPSVYPL GQGTLVVSSASTKGPSVFPL GQGTLIVSSASTKGPSVFPL GQGTLIVSSASTTPSVYPL GQGTLIVSSASTTPSVYPL GQGTLVVSSASTTPSVYPL GQGTLIVSSASTTPSVYPL GQGTLIVSSASTTPPSVYPL GQGTLIVSSASTTPPSVYPL GQGTLVVSSASTKGPSVPPL GQGTVVVSSASTKGPSVPPL GQGTVVVSSASTKGPSVPPL GQGTVVVSSASTKGPSVPPL	APGS APGG APCS APSS APVC APGS APGS APGS APGS APGS APGS APGS APGS	AAQTNSMVTLG 3GGGGAWVTL KSTSGTAAL KSTSGTAAL ADTTGSVTL AAQTNSMVTL AAQTNSMVTL AAQTNSMVTL AAQTNSMVTL AAQTNSMVTL AAQTNSMVTL KSTSGTAAL KSTSGTAAL KSTSGTAAL	LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV	TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA	/LQS-DLYTLSSSVTVPSS: /LQS-DLYTLSSSVTVPSS: /LQSSGLYSLSSVVTVPSS: /LQSSGLYSLSSVVTVPSS: /LQS-DLYTLSSSVTVPSS: /LQS-CLYTMSSSVTVPSS: /LQS-CLYTMSSSVTVPSS: /LQS-DLYTLSSVTVPSS: /LQS-DLYTLSSVTVPSS: /LQS-DLYTLSSVTVPSS: /LQS-DLYTLSSVTVPSS: /LQS-CLYTMSSSVTVPSS: /LQS-CLYTLSSSVTVPSS: /LQSSGLYSLSSVVTVPSS:	TWPSETVTONVAHE PWPSETVTONVAHE SLGTATYT NVDHK SLGTQTY I ONVNHK TWPSQSITONVAHE PRPSETVTONVAHE PRPSETVTONVAHE TWPSQTVTONVAHE TWPSQTVTONVAHE SLGTQTYI ONVAHE SLGTQTYI ONVNHK SLGTQTYI ONVNHK	ASSTKVDKKIVPRDC ASSTKVDKKIVPRDC VPSNTKVDKKVEPKSCDKT VPSNTKVDKKIEPR ASSTKVDKKIVPR ASSTKVDKKIVPROC ASSTTVDKKIVF ASSTKVDKKIVPROCGGI ASSTKVDKKIVPROCGGI ASSTKVDKKIVPROCHHF (PSNTKVDKKVEPKSCDKT)	225 215 219 219 218 218 217 7C 223 215 219 HHH- 226 217 'SPLF 228	- 5 3 5 5 9 9 3 5 7 3 5 9 5 7 3 3
1FNS_HC 1HOD_HC 1IQD_HC 1JPS_HC 1KB5_HC 1LK3_HC 1MLC_HC 1OSP_HC 1UJ3_HC 1YQV_HC 1YQV_HC 1ZTX_HC 2B2X_HC 2CMR_HC 2PD6_HC 2NXY_HC	GQGTSVTVSSAKTTPPSVYPL GQGTSVTVSSAKTTPPSVYPL GQGTWTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTTAPSVYPL GQGTLVTVSAKTTPPSVYPL GQGTLVTVSAKTTPPSVYPL GQGTLTVSSASTTPPSVYPL GQGTLTLVSSAKTTPPSVYPL GQGTLTVVSSASTTPSVYPL GQGTLVTVSSASTKGPSVFPL GQGTTVTVSSASTKGPSVFPL GQGTTVTVSSASTKGPSVFPL GQGTTVTVSSASTKGPSVFPL	APGS APGG APCS APVC APGT APGS APGC APGS APGC APGS APGS APGS APGS APGS	AAQTNSMVTLG GGGGQAWTIL KSTSGTAAL KSTSGGTAAL GDTTGSSVTLG AAQTNSMVTLG AAQTNSMVTLG GDTGSSVTLG AAQTNSMVTLG GDTGSSVTLG AAQTNSMVTLG RSTSGGTAAL KSTSGGTAAL KSTSGGTAAL	CLVKGYFPEPVTV CLVKGYFPEPVTV CLVKGYFPEPVTV LVKGYFPEPVTV CLVKGYFPEPVTV CLVKGYFPEPVTV CLVKGYFPEPVTV CLVKGYFPEPVTV CLVKGYFPEPVTV CLVKGYFPEPVTV CLVKGYFPEPVTV CLVKGYFPEPVTV CLVKGYFPEPVTV CLVKGYFPEPVTV CLVKGYFPEPVTV CLVKGYFPEPVTV CLVKGYFPEPVTV	TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA	/LQS-DLYTLSSSVTVPSS /LQS-DLYTLSSSVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS /LQS-DLYTLSSSVTVPSS /LQS-DLYTLSSSVTVPSS /LQS-GLYTMSSSVTVPSS /LQS-GLYTMSSSVTVPSS /LQS-GLYTLSSSVTVPSS /LQS-GLYTLSSSVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS	TWPSETVTONVAHE PWPSETVTONVAHE SLGTQTYIONVAHE SLGTQTYIONVAHE TWPSQTVTONVAHE TWPSQTVTONVAHE SLGTKTYTONVAHE TWPSQTVTONVAHE TWPSETVTONVAHE SLGTQTYIONVAHE SLGTXINO	ASSTKVDKKIVPRDC ASSTKVDKKIVPRDC PASTKVDKKVEPKSCDKT ASSTKVDKKIVPR ASSTKVDKKIVPR ASSTKVDKKIVPRC ASSTVDKKLVPRCGGL ASSTKVDKKIVPRNCGGL ASSTKVDKKIVPRCGHH (PSNTKVDKKVEPKSCDKT ASSTKVDKKIAAA	225 215 'HT 225 219 218 218 218 213 215 215 217 'SPLF 228 217 'SPLF 228 213	- 5 3 5 5 3 5 5 3 5 7 3 5 9 5 7 3 3 9 5 7 3 3 9
1FNS-HC 1HOD-HC 1JQD-HC 1JPS-HC 1KB5-HC 1MLC-HC 1MLC-HC 10SP-HC 1UJ3-HC 1WEJ-HC 1YQV-HC 1ZTX_HC 2CMR-HC 2DD8-HC 2DN8-HC 2Q8B-HC	GQGTSVTVSSAKTTPPSVYPL GQGTSVTVSSASTKGPSVPPL GQGTWVVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTTAPSVYPL GQGTLVTVSSASTKGPSVPPL GQGTLTVSSASTKGPSVPPL GQGTLTVSSAKTTPPSVVPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTTVTVSSASTKGPSVPPL GQGTTVTVSSASTKGPSVPPL GQGTTVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL	APGS APGG APCS APCS APCC APGT APGS APGS APGS APGS APGS APGS APGS APS APS APS APS	AAQTNSMVTLG GGGGGAWVTLG SGTSGSTAALG KSTSGTAALG GDTTGSSVTLA ALKSNSMVTLG GDTTGSSVTLG GDTTGSSVTLG GDTTGSSVTLG AAQTNSMVTLG GDTTGSSVTLG AAQTNSMVTLG CSTSGGTAALG KSTSGGTAALG KSTSGGTAALG KSTSGGTAALG GDTTGSSVTLG	LVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV SLVKGYFPEPVTV	TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA SWNSGALTSGVHTFPA SWNSGALTSGVHTFPA TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA SWNSGALTSGVHTFPA SWNSGALTSGVHTFPA SWNSGALTSGVHTFPA TWNSGSLSSGVHTFPA	/LQS-DLYTLSSSVTVPSS' /LQS-DLYTLSSSVTVPSS: /LQSSGLYSLSSVVTVPSS: /LQSSGLYSLSSVVTVPSS' /LQS-GLYTLSSSVTVPSS' /LQS-GLYTLSSSVTVPSS' /LQS-GLYTMSSSVTVPSS' /LQS-DLYTLSSSVTVPSS' /LQS-GLYTMSSSVTVPSS' /LQS-GLYTMSSSVTVPSS' /LQSSGLYSLSSVVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS'	TWPSETVTONVAHE PWPSETVTONVAHE SLGJTYTINVDHE SLGJTYTINVDHE TWPSQITONVAHE TWPSQTVTONVAHE TWPSQTVTONVAHE TWPSQTVTONVAHE TWPSQTVTONVAHE TWPSETVTONVAHE SLGJQTTIONVAHE SLGJQTIONVAHE SLGJQTTIONVAHE SLGJQTTIONVAHE SLGJQTTIONVAHE SLGJQTTIONVAHE SLGJQTTIONVAHE SLGJQTTIONVAHE SLGJQTTIONVAHE SLGJQTTIONVAHE SLGJQTTIONVAHE SLGJQTTIONVAHE SLGJQTTIONVAHE SLGJQTTIONVAHE SLGJQTTIONVAHE SLGJQTTIONVAHE SLGJQTTIONVAHE SL	ASSTKVDKKIVPRDC ASSTKVDKKIVPRDC CPSNTKVDKKVEPKSCDKT ASSTKVDKKIVPR ASSTKVDKKIVPRDC ASSTKVDKKIVPRDC ASSTKVDKKIVPRCGGI ASSTKVDKKIVPRCGGI ASSTKVDKKILD ASSTKVDKKIVPRCGGI ASSTKVDKKIVPRCGHI ASSTKVDKKIVPRCHH CPSNTKVDKKVEPKSCDKT ASSTKVDKKIAA CPSNTKVDKKVEPK	225 213 YHT 225 YHT 225 219 218 218 218 YC 223 219 YSPLF 228 217 YSPLF 228 213 YSPLF 228 212	- 535599337359573390
1FNS-HC HOD-HC IIQD-HC IIQD-HC ILK3-HC ILK3-HC ILK3-HC IUJ3-HC IUJ3-HC IUJ3-HC IUJ3-HC IZTX-HC 2D28-HC 2D28-HC 2FD6-HC 2Q08-HC 2Q08-HC	GQGTSVTVSSAKTTPPSVYPL GQGTSVTVSSAKTTPPSVYPL GQGTLVTVSSASTKGPSVFPL GQGTL/TVSSASTKGPSVFPL GQGTL/TVSSASTTPSVYPL GQGTL/TVSSASTTPSVYPL GQGTL/TVSSASTTPSVYPL GQGTL/TVSSASTTPSVYPL GQGTL/TVSSASTTPSVYPL GQGTL/TVSSASTTPSVYPL GQGTL/TVSSASTKGPSVPPL GQGTL/TVSSASTKGPSVPPL GQGTL/TVSSASTKGPSVPPL GQGTL/TVSSASTKGPSVPPL GQGTL/TVSSASTKGPSVPPL GQGTL/TVSSASTKGPSVPPL GQGTL/TVSSASTKGPSVPPL	APGS APGG APCS APCS APCC APGT APGS APGS APGS APGS APGS APGS APS APS APS APS APS	AAQTNSMVTLG GGGGAMVTLG SGTSGTAALG STSESTAALG STSESTAALG ALKSNSMVTLG AAQTNSMVTLG AAQTNSMVTLG AAQTNSMVTLG AAQTNSMVTLG AAQTNSMVTLG ASTSGTAALG KSTSGGTAALG SGTSGSVTLG SGTGSSVTLG SGTSGSVTLG	LVKGYFPEPVTV LVKGYFPEPVTV	TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSSVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA	/LQS-DLYTLSSSVTVPSS: /LQS-DLYTLSSSVTVPSS: /LQSSGLYSLSSVVTVPSS: /LQSSGLYSLSSVVTVPSS: /LQS-DLYTLSSSVTVPSS: /LQS-DLYTLSSVTVPSS: /LQS-CLYTMSSSVTVPSS: /LQS-DLYTLSSVTVPSS: /LQS-DLYTLSSVTVPSS: /LQS-DLYTLSSVTVPSS: /LQS-DLYTLSSVTVPSS: /LQSSGLYSLSSVVTVPSS: /LQSSGLYSLSSVVTVPSS: /LQSSGLYSLSSVVTVPSS: /LQSSGLYSLSSVVTVPSS: /LQSSGLYSLSSVVTVPSS:	TWPSETVTONVAHE PWPSETVTONVAHE SLGTQTYIONVAHE SLGTQTYIONVAHE PRPSETVTONVAHE PRPSETVTONVAHE TWPSQTVTONVAHE TWPSQTVTONVAHE SLGTQTYIONVAHE	ASSTKVDKKIVPRDC ASSTKVDKKIVPRDC CPSNTKVDKKVEPKSCDKT ASSTKVDKKIPR ASSTKVDKKIVPR ASSTKVDKKIVPR ASSTTVDKKIP ASSTKVDKKIVPRNCGGI ASSTKVDKKIVPRNCGGI ASSTKVDKKIP ASSTKVDKKIP ASSTKVDKKIP ASSTKVDKKIPA CPSNTKVDKKVEPKSCDKT ASSTVDKKLE	225 223 215 219 219 218 217 217 217 217 217 219 210 	- 5 3 5 5 9 9 3 3 7 3 5 9 5 7 3 3 9 0 0
1FNS-HC 1HOD-HC 1JQD-HC 1JPS-HC 1KB5-HC 1MLC-HC 1MLC-HC 10SP-HC 1UJ3-HC 1WEJ-HC 1YQV-HC 1ZTX_HC 2CMR-HC 2DD8-HC 2DN8-HC 2Q8B-HC	GQGTSVTVSSAKTTPPSVYPL GQGTSVTVSSASTKGPSVPPL GQGTWVVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTTAPSVYPL GQGTLVTVSSASTKGPSVPPL GQGTLTVSSASTKGPSVPPL GQGTLTVSSAKTTPPSVVPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTTVTVSSASTKGPSVPPL GQGTTVTVSSASTKGPSVPPL GQGTTVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL	APGS APGG APCS APSS APVC APGT APGS APGS APGS APGS APGS APGS APGS APGS	AAQTNSMVTLG GGGGGAWTIL KSTSGGTAAL KSTSGGTAAL GDTTGSSVTLG AAQTNSMVTLG AAQTNSMVTLG STSGSSVTLG KSTSGGTAAL KSTSGGTAAL CSTSGGTAAL CSTSGGTAAL CSTSGGTAAL CSTSGGTAAL CSTSGGTAAL CSTSGGTAAL	LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPEPVTV LVKGYFPESVTV LVKGYFPESVTV LVKGYFPESVTV LVKGYFPESVTV LVKGYFPESVTV LVKGYFPEPVTV LVKGYFPESVTV	TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA	/LQS-DLYTLSSSVTVPSS /LQS-DLYTLSSSVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS /LQS-DLYTLSSSVTVPSS /LQS-DLYTLSSSVTVPSS /LQS-GLYTMSSSVTVPSS /LQS-GLYTMSSSVTVPSS /LQS-GLYTMSSSVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS	TWPSETVTONVAHE PWPSETVTONVAHE SLGTQTYIONVAHE SLGTQTYIONVAHE TWPSQTVTONVAHE PRPSETVTONVAHE SLGTVTTONVAHE TWPSQTVTONVAHE TWPSQTVTONVAHE SLGTQTYIONVAHE SLGTVINA SLGTQTYIONVAHE SLGTQTYIONVAHE SLGTQTYIONVAHE SLGTVINA	ASSTKVDKKIVPRDC ASSTKVDKKIVPRDC VPSNTKVDKKVEPKSCDKT ASSTKVDKKIVPR ASSTKVDKKIVPR ASSTKVDKKIVPR ASSTKVDKKIVPROC ASSTKVDKKIVPRNCGGI ASSTKVDKKIPROCHH (PSNTKVDKKVEPKSCDKT (PSNTKVDKKVEPKS ASSTTVDKKIPROCHH (PSNTKVDKKVEPKSCDT)	225 213 215 219 218 218 217 215 217 217 217 217 217 217 210 210 210 229 210 220	533559933735957339000
1FNS-HC 1HOD-HC 1JQD-HC 1JPS-HC 1KB5-HC 1LK3-HC 1USP-HC 1UJ3-HC 1VQV-HC 1ZTX-HC 2DD8-HC 2FD6-HC 2FD6-HC 2QN-HC 2Q0N-HC 2R0L-HC	GQGTSVTVSSAKTTPPSVYPL GQGTSVTVSSAKTPPSVYPL GQGTWTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLIVSSASTTAPSVYPL GQGTLIVSSASTTPPSVYPL GQGTLIVSSAKTTPPSVYPL GQGTLIVSSAKTTPPSVYPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL	APGSJ APGGAPCSJ APSSJ APVCTJ APGSJ APGCI APGSJ APGCSJ APGSJ APGSJ APGSJ APGSJ APGSJ APGSSJ APGSSJ APGSSJ APSSJ APSSJ APSSJ	AAQTNSMVTLG GGGGQMVTLA SSTSESTAALA SSTSESTAALA SDTGSSVTLA ALKSNSMVTLA GDTTGSSVTLA AAQTNSMVTLA GDTTGSSVTLA AAQTNSMVTLA AAQTNSMVTLA AAQTNSMVTLA CSTSGGTAALA SSTSGGTAALA SSTSGGTAALA SSTSGGTAALA SSTSGGTAALA SSTSGGTAALA SSTSGGTAALA SSTSGGTAALA SSTSGGTAALA SSTSGGTAALA SSTSGGTAALA SSTSGGTAALA SSTSGGTAALA SSTSGGTAALA SSTSGGTAALA SSTSGGTAALA SSTSGGTAALA SSTSGGTAALA SSTSGGTAALA SSTSGGTAALA	LVKGYFPEPVTV LVKGYFPEPVTV	TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA	/LQS-DLYTLSSSVTVPSS /LQS-DLYTLSSSVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS /LQS-GLYTLSSSVTVPSS /LQS-GLYTLSSVTVPSS /LQS-GLYTMSSSVTVPSS /LQS-GLYTMSSSVTVPSS /LQS-GLYTMSSSVTVPSS /LQS-GLYTLSSVTVPSS /LQS-GLYSLSSVVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS	TWPSETVTONVAHE PWPSETVTONVAHE SLGTQTYIONVAHE SLGTQTYIONVAHE TWPSQUTTONVAHE TWPSQTVTONVAHE TWPSQTVTONVAHE TWPSQTVTONVAHE TWPSETVTONVAHE SLGTQTYIONVAHE	ASSTKVDKKIVPRDC ASSTKVDKKIVPRDC ASSTKVDKKVEPKSCDKT ASSTKVDKKIVPR ASSTKVDKKIVPRC ASSTKVDKKIVPRCGE ASSTKVDKKIVPRCGE ASSTKVDKKIVPRCGE ASSTKVDKKIVPRCGE ASSTKVDKKIVPRCGE ASSTKVDKKIVPRCHE ASSTKVDKKIPS ASSTKVDKKIAAA ASSTKVDKKIAAA ASSTTVDKKE ASSTKVDKKIAAA ASSTTVDKKE ASSTKVDKKIAAA ASSTVDKKE ASSTKVDKE ASSTKV	225 215 YHT 225 YHT 225 219 218 218 218 217 YSPLF 228 217 YSPLF 228 217 YSPLF 228 210 YSPLF 228 210 YSPLF 228 210 YSPLF 228 210 YSPLF 228 210 YSPLF 228 210 YSPLF 228 210 YSPLF 228 210 YSPLF 228 210 YSPLF 228 	- 5 3 5 5 9 9 3 7 3 5 9 5 7 3 3 9 0 0 1
1FNS <sup>-</sup> HC 1HQD <sup>-</sup> HC 1JQP <sub>3</sub> HC 1KB5 <sup>-</sup> HC 1KK3 <sup>-</sup> HC 1MLC <sup>-</sup> HC 1OSP <sup>-</sup> HC 1UJ3 <sup>-</sup> HC 1VQV <sup>-</sup> HC 1ZTX <sup>-</sup> HC 2DD8 <sup>-</sup> HC 2CMR <sup>-</sup> HC 2DD8 <sup>-</sup> HC 2DD8 <sup>-</sup> HC 2Q8 <sup>-</sup> BHC 2Q8 <sup>-</sup> BHC 2Q8 <sup>-</sup> HC 2Q0 <sup>-</sup> HC 2Q0 <sup>-</sup> HC 2VD2 <sup>-</sup> HC	GQGTSVTVSSAKTTPPSVYPL GQGTSVTVSSAKTTPPSVYPL GQGTWTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTTAPSVYPL GQGTLVTVSSASTTPPSVYPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL	APGSJ APGG APCSS APSS APVCC APGZ APGZ APGZ APGZ APGZ APGZ APGS APGC APSS APGC APSS APGC APSS APVC APSS	AAQTNSMVTLG GGGGAMVTLG SGTSSTAALG SGTSSTAALG SGTSSTAALG SGTSSTAALG ALKSNSMVTLG AAQTNSMVTLG AAQTNSMVTLG AAQTNSMVTLG AAQTNSMVTLG AAQTNSMVTLG RSTSGTAALG KSTSGTAALG KSTSGTAALG KSTSGTAALG SGTSGSVTLG SGTGGSVTLG SGTGGSVTLG SGTGGSVTLG SGTGGSVTLG SGTGGSVTLG SGTGGSVTLG SGTGGSVTLG SGTGGSVTLG SGTGGSVTLG SGTGGSVTLG SGTGGSVTLG SGTGGSVTLG SGTGGSVTLG SGTGGSVTLG SGTGGSVTLG SGTGGSVTLG	2 LVKGYFPEPVTV 3 LVKGYFPEPVTV 3 LVKGYFPEPVTV 3 LVKGYFPEPVTV 4 LVKGYFPEPVTV 4 LVKGYFPEPVTV 5 LVKGYFPEPVTV	TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA SWNSGALTSGVHTFPA SWNSGALTSGVHTFPA TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA SWNSGALTSGVHTFPA SWNSGALTSGVHTFPA SWNSGALTSGVHTFPA SWNSGALTSGVHTFPA SWNSGALTSGVHTFPA	/LQS-DLYTLSSSVTVPSS: /LQS-DLYTLSSSVTVPSS: /LQSSGLYSLSSVTVPSS: /LQSSGLYSLSSVTVPSS: /LQS-DLYTLSSSVTVPSS: /LQS-DLYTLSSVTVPSS: /LQS-CLYTMSSSVTVPSS: /LQS-DLYTLSSVTVPSS: /LQS-DLYTLSSVTVPSS: /LQS-DLYTLSSVTVPSS: /LQS-DLYTLSSVTVPSS: /LQSSGLYSLSSVTVPSS: /LQSSGLYSLSSVTVPSS: /LQSSGLYSLSSVTVPSS: /LQSSGLYSLSSVTVPSS: /LQSSGLYSLSSVTVPSS: /LQSSGLYSLSSVTVPSS: /LQSSGLYSLSSVTVPSS: /LQSSGLYSLSSVTVPSS: /LQSSGLYSLSSVTVPSS: /LQSSGLYSLSSVTVPSS:	TWPSETVT NVAHE PWPSETVT NVAHE SLGTQTY I NVDHK SLGTQTY I NVNHK SLGTQTY I NVNHK PRPSETVT NVAHE PRPSETVT NVAHE TWPSQTVT SVAHE TWPSQTVT NVAHE SLGTQTY I NVNHK SLGTQTY I NVNHK	ASSTKVDKKIVPRDC ASSTKVDKKIVPRDC CPSNTKVDKKVEPKSCDKT ASSTKVDKKIPR PASSTKVDKKIVPR PASSTKVDKKIVPR PASSTKVDKKIVPRCC PASSTKVDKKIVPRCGGI ASSTKVDKKIPS PASSTKVDKKIPS PASSTKVDKKIPRCCHH CPSNTKVDKKVEPKSCDKT CPSNTKVDKKVEPKSCCKT CPSNTKVDKKVEPKSCCKT CPSNTKVDKKVEPKSCC PASSTKVDKKIPRC PASSTKVDKKIPRSC	225 215 215 219 219 218 218 217 217 217 217 210 210 210 210 210 220 220 220 220 220 220 220 220 220 220 220 220 220 220 	- 5 3 5 5 9 9 3 3 7 3 5 9 5 7 3 3 9 0 0 L 3
1FNS-HC HOD_HC 1IQD-HC 1IQD-HC 1IQDS-HC 1KB5-HC 1MLC-HC 1003P-HC 1UJ3-HC 1UJ3-HC 1UJ3-HC 1VEJ-HC 1ZTX-HC 2DD8-HC 2DD8-HC 2QQN-HC 2QDR-HC 2VDR-HC 2VDR-HC	GQGTSVTVSSAKTTPPSVYPL GQGTSVTVSSAKTTPPSVYPL GQGTLVTVSSASTKGPSVFPL GQGTL/TVSSASTKGPSVFPL GQGTL/TVSSASTRSPSVPL GQGTL/TVSSASTTPSVYPL GQGTL/TVSSASTTPSVYPL GQGTL/TVSSASTTPSVYPL GQGTL/TVSSASTTPSVYPL GQGTL/TVSSASTTPSVYPL GQGTLVTVSSASTKGPSVPPL GQGTTVTVSSASTKGPSVPPL GQGTL/TVSSASTKGPSVPPL GQGTL/TVSSASTKGPSVPPL GQGTL/TVSSASTKGPSVPPL GQGTL/TVSSASTKGPSVPPL GQGTL/TVSSASTKGPSVPPL GQGTL/TVSSASTKGPSVPPL GQGTL/TVSSASTKGPSVPPL GQGTL/TVSSASTKGPSVPPL GQGTL/TVSSASTKGPSVPPL GQGTL/TVSSASTKGPSVPPL	APGSJ APGGAPCS: APSC: APGC: APGC: APGC: APGC: APGSJ APGC: APGSJ APGSS APGSS APGC: APSSS APGC: APSSS APSCC APSSS APSCC APSSS APCSJ	AAQTNSMVTLG GGGGAMVTLG SGTSSTAALG SGTSSTAALG SGTSSTAALG SGTSSTAALG SGTSGSVTLG AAQTNSMVTLG SGTSGSSVTLG SGTSGGTAALG	LVKGYFPEPVTV LVKGYFPEPVTV	TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALSGVHTFPA 'SWNSGALSGVHTFPA 'SWNSGALSGVHTFPA 'SWNSGALSGVHTFPA 'SWNSGALSGVHTFPA 'SWNSGALSGVHTFPA 'SWNSGALSGVHTFPA	/LQS-DLYTLSSSVTVPSS /LQS-DLYTLSSSVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS /LQS-DLYTLSSSVTVPSS /LQS-DLYTLSSVTVPSS /LQS-GLYTMSSSVTVPSS /LQS-GLYTMSSSVTVPSS /LQS-GLYTMSSSVTVPSS /LQS-GLYTMSSSVTVPSS /LQS-GLYTMSSSVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS	TWPSETVTONVAHE PWPSETVTONVAHE SLGTQTYIONVAHE SLGTQTYIONVAHE PWPSETVTONVAHE PWPSETVTONVAHE SLGTKTYTONVAHE TWPSETVTONVAHE TWPSETVTONVAHE SLGTQTYIONVAHE	ASSTKVDKKIVPRDC ASSTKVDKKIVPRDC YASTKVDKKVEPKSCDKT ASSTKVDKKIVPR ASSTKVDKKIVPR YASSTKVDKKIVPRC YASSTKVDKKIVPRDC YASSTKVDKKIVPRDC YASSTKVDKKIPRCGGI YASSTKVDKKIPRCGGI YASSTKVDKKIPRC YASSTKVDKKIPRSCHH (PSNTKVDKKIPRSCHT YASSTKVDKKIPRSCCHT (PSNTKVDKKVEPKSCDKT YASSTKVDKKIPRSCCHT (PSNTKVDKKVEPKSCCHT) YASSTKVDKKIPRG YASSTKVDKKIPRG YASSTKVDKKIPRG	225 213 215 219 218 218 218 215 217 217 217 219 210 210 210 220 221 220 220 	- 5 5 3 5 5 9 9 3 3 7 3 5 9 5 7 3 3 9 5 7 5 9 5 7 3 3 9 0 0 0 1 3 0 0 0 1 3 0 0 0 1 3 0 0 0 0 1 0 0 0 0
1FNS-HC 1HOD-HC 1JQD-HC 1JPS-HC 1K85-HC 1LK3-HC 1USJ-HC 1UJ3-HC 1VQV-HC 1YQV-HC 2D2K-HC 2CMR-HC 2DB8-HC 2MXY-HC 2Q8B-HC 2Q0H-C 2Q0L-HC 2C0L-HC 2D0A-HC	GQGTSVTVSSAKTPPSVYPL GQGTSVTVSSAKTPPSVYPL GQGTMVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTRPSVYPL GQGTLIVSSASTTPSVYPL GQGTLVTVSASTTPSVYPL GQGTLVVSSASTTPSVYPL GQGTLVVSSASTTPSVYPL GQGTLVTVSSASTRGPSVYPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSASTKGPSVFPL GQGTLVTVSASTKGPSVFPL GQGTLVTVSASTKGPSVFPL GQGTLVTVSASTKGPSVFPL	APGSJ APGGAPCSS APSS APSS APGT APGSJ APGCAPCSS APGCAPCSS APGCAPCSS APGSS APSSS APSSS APSSS APSSS APSSS APSSS	AAQTNSMVTLG GGGGQMVTLA SSTSESTAALA KSTSGGTAALA GDTTGSSVTLA AAQTNSMVTLA AAQTNSMVTLA GDTTGSSVTLA AAQTNSMVTLA GDTTGSSVTLA KSTSGGTAALA GDTTGSSVTLA SSTSGTAALA GDTTGSSVTLA GSTSGTAALA	LVKGYFPEPVTV SLVKGYFPEPVTV	TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA	/LQS-DLYTLSSSVTVPSS /LQS-DLYTLSSSVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS /LQS-GLYTLSSSVTVPSS /LQS-GLYTLSSSVTVPSS /LQS-GLYTMSSSVTVPSS /LQS-GLYTMSSSVTVPSS /LQS-GLYTMSSSVTVPSS /LQS-GLYTMSSSVTVPSS /LQS-GLYTMSSSVTVPSS /LQS-GLYTMSSSVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS	TWPSETVTONVAHE PWPSETVTONVAHE SLGTQTYIONVAHE SLGTQTYIONVAHE TWPSQITONVAHE TWPSQTVTONVAHE TWPSQTVTONVAHE SLGTKTYTONVAHE TWPSCTVTONVAHE SLGTQTYIONVAHE	ASSTKVDKKIVPRDC ASSTKVDKKIVPRDC ASSTKVDKKVEPKSCDKT ASSTKVDKKIVPR ASSTKVDKKIVPRC ASSTKVDKKIVPRC ASSTKVDKKIVPRNCGGE ASSTKVDKKIVPRNCGGE ASSTKVDKKIVPRNCGGE ASSTKVDKKIVPRCHHE ASSTKVDKKIVPRSCDKT ASSTVDKKVEPKSCDKT ASSTVDKKVEPKSCD ASSTVDKKVEPKSC	225 215 'HT 225 218 218 218 218 217 'SPLF 228 217 'SPLF 228 210 'H 220 221 'HT 223 210 'HT 231	- 5 5 3 5 5 9 9 3 3 7 3 5 9 5 7 3 3 9 0 0 0 1 3 0 1 2 0 1 2 0 1 2 0 1 2 0 0 1 1 2 0 0 1 1 1 0 1 0
1FNS-HC 1HOD-HC 1IQD-HC 1JPS-HC 1KB5-HC 1LK3-HC 1UJ3-HC 1UJ3-HC 1VQV-HC 1ZTX-HC 2DD8-HC 2DD8-HC 2DD8-HC 2Q8B-HC 2Q8B-HC 2Q0N-HC 2Q0LHC 2VDR-HC 3D9A-HC 3D9A-HC	GQGTSVTVSSAKTTPPSVYPL GQGTSVTVSSAKTTPPSVYPL GQGTWTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTTAPSVYPL GQGTLITVSSASTTAPSVYPL GQGTLVTVSSASTTPPSVYPL GQGTLVTVSSAKTTPPSVYPL GQGTLVTVSSASTKGPSVFPL	APGSJ APGG APCSJ APSSJ APSCJ APGCJ APGSJ APGSJ APGSJ APGSJ APGSJ APGSSJ APSSSJ APSSSJ APSSJ APSSJ APSSJ APSSJ APSSJ APSSJ	AAQTNSMVTLG GGGGAMVTLG SGTSESTAALG SGTSESTAALG SGTSESTAALG SGTSESTAALG AAQTNSMVTLG AAQTNSMVTLG AAQTNSMVTLG AAQTNSMVTLG AAQTNSMVTLG RSTSEGTAALG KSTSGGTAALG KSTSGGTAALG KSTSGGTAALG KSTSGGTAALG KSTSGGTAALG KSTSGGTAALG KSTSGGTAALG KSTSGGTAALG KSTSGGTAALG KSTSGGTAALG KSTSGGTAALG KSTSGGTAALG KSTSGGTAALG KSTSGGTAALG	2 LVKGYFPEPVTV 3 LVKGYFPEPVTV 3 LVKGYFPEPVTV 3 LVKGYFPEPVTV 3 LVKGYFPEPVTV 3 LVKGYFPEPVTV 3 LVKGYFPEPVTV 3 LVKGYFPEPVTV 4 LVKGYFPEPVTV 4 LVKGYFPEPVTV 5 LVKGYFPEPVTV 5 LVKGYFPEPVTV 4 LVKGYFPEPVTV 5 LVKGYFPEPVTV	TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA SWNSGALTSGVHTFPA SWNSGALTSGVHTFPA TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA SWNSGALTSGVHTFPA SWNSGALTSGVHTFPA SWNSGALTSGVHTFPA SWNSGALTSGVHTFPA SWNSGALTSGVHTFPA SWNSGSLSSGVHTFPA SWNSGSLSSGVHTFPA	/LQS-DLYTLSSSVTVPSS: /LQS-DLYTLSSSVTVPSS: /LQSSGLYSLSSVTVPSS: /LQSSGLYSLSSVTVPSS: /LQS-CLYTLSSSVTVPSS: /LQS-CLYTLSSVTVPSS: /LQS-CLYTMSSSVTVPSS: /LQS-DLYTLSSVTVPSS: /LQS-DLYTLSSVTVPSS: /LQS-DLYTLSSVTVPSS: /LQSSGLYSLSSVTVPSS: /LQSSGLYSLSSVTVPSS: /LQSSGLYSLSSVTVPSS: /LQSSGLYSLSSVTVPSS: /LQSSGLYSLSSVTVPSS: /LQSSGLYSLSSVTVPSS: /LQSSGLYSLSSVTVPSS: /LQSSGLYSLSSVTVPSS: /LQSSGLYSLSSVTVPSS: /LQSSGLYSLSSVTVPSS: /LQSSGLYSLSSVTVPSS: /LQSSGLYSLSSVTVPSS: /LQSSGLYSLSSVTVPSS: /LQSSGLYSLSSVTVPSS: /LQSSGLYSLSSVTVPSS:	TWPSETVT NVAHE PWPSETVT NVAHE SLGTQTY I NVDHK SLGTQTY I NVDHK SLGTQTY I NVNHK PRPSETVT NVAHE PRPSETVT NVAHE TWPSQTVT SVAHE TWPSQTVT NVAHE TWPSQTVT NVAHE SLGTQTY I NVNHK SLGTQTY I NVNHK	ASSTKVDKKIVPRDC ASSTKVDKKVPRDC PASSTKVDKKVEPKSCDKT ASSTKVDKKIVPR PASSTKVDKKIVPR PASSTKVDKKIVPR PASSTKVDKKIVPR PASSTKVDKKIVPRCGGI PASSTKVDKKIPRC PASSTKVDKKIPRC	225 223 215 219 219 218 218 217 217 210 210 210 210 220 210 220 210 220 210 220	- 5 5 3 5 5 9 9 3 3 7 3 5 9 5 7 3 3 9 0 0 0 1 3 0 1 2 0 1 2 0 1 2 0 1 2 0 0 1 1 2 0 0 1 1 1 0 1 0
1FNS-HC HOD_HC 1IQD-HC 1IQD-HC 1IQD-HC 1IVB5-HC 1MK5-HC 1MLC-HC 1003P-HC 1UJ3-HC 1UJ3-HC 1VEJ-HC 1ZTX-HC 2DD8-HC 2DD8-HC 2QQN-HC 2QVD-HC 2VDR-HC 3D85-HC 3D81-HC 1SC1-HC	GQGTSVTVSSAKTTPPSVYPL GQGTSVTVSSAKTTPPSVYPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLTVSSASTTPSVYPL GQGTLVTVSSASTTPSVYPL GQGTLVTVSSASTTPSVYPL GQGTLVTVSSASTTPSVYPL GQGTLVTVSSASTTPSVYPL GQGTLVTVSSASTTPSVYPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGLVTVSSASTKGPSVPPL	APGSJ APGG APCSJ APSSJ APSCJ APGCJ APGTJ APGSJ APGCJ APGSJ APGCJ APGSJ APGCSJ APGSJ APSSJ APSSJ APSSJ APSSJ APSSJ APSSSJ APSSJ APSSJ	AAQTNSMVTLG 3GGGGAWTIL KSTSGTAAL KSTSGTAAL STSGTAAL GDTTGSSVTL AAQTNSMVTL AAQTNSMVTL AAQTNSMVTL GDTTGSSVTL GDTTGSSVTL GDTTGSSVTL GDTTGSSVTL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL	LVKGYFPEPVTV LVKGYFPEPVTV	TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALSGVHTFPA 'SWNSGALSGVHTFPA 'SWNSGALSGVHTFPA 'SWNSGALSGVHTFPA 'SWNSGALSGVHTFPA 'SWNSGALSGVHTFPA 'SWNSGALSGVHTFPA 'SWNSGALSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA	/LQS-DLYTLSSSVTVPSS /LQS-DLYTLSSSVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVVTVPSS /LQS-DLYTLSSSVTVPSS /LQS-DLYTLSSVTVPSS /LQS-DLYTLSSVTVPSS /LQS-GLYTMSSSVTVPSS /LQS-GLYTMSSSVTVPSS /LQS-GLYTLSSVTVPSS /LQS-GLYTLSSSVTVPSS /LQS-GLYTLSSSVTVPSS /LQSSGLYSLSSVVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS /LQSSGLYSLSSVTVPSS	TWPSETVT NVAHE PWPSETVT NVAHE SLGTQTY I NVDHK SLGTQTY I NVNHK SLGTQTY I NVNHK PRPSETVT NVAHE PRPSETVT NVAHE SLGTKYT NVAHE TWPSQTVT SVAHE TWPSQTVT SVAHE SLGTQTY I NVNHK SLGTQTY I NVNHK	ASSTKVDKKIVPRDC ASSTKVDKKIVPRDC QSNTKVDKKVEPKSCDKT ASSTKVDKKIVPR ASSTKVDKKIVPR ASSTKVDKKIVPR ASSTKVDKKIVPRCGGI ASSTKVDKKIVPRCGGI ASSTKVDKKIVPRCGGI ASSTKVDKKIVPRCGGI ASSTKVDKKIVPRCGGI ASSTKVDKKIVPRCCHH (PSNTKVDKKVEPKSCDKT (PSNTKVDKKVEPKSCDKT ASSTKVDKKVEPKSCDKT ASSTKVDKKVEPKSCDKT ASSTKVDKKVEPKSCDKT ASSTKVDKKVEPSCDKT ASS	225 213 215 219 218 218 218 217 213 217 213 217 210 210 210 210 210 210 210 210	- 5 5 3 5 5 9 9 3 3 7 3 5 9 5 7 3 3 9 0 0 0 L 3 0 L 0 0 0 L 0 0 0
1FNS-HC 1HOD-HC 1IQD-HC 1JPS-HC 1K85-HC 1LK3-HC 1UJ3-HC 1UJ3-HC 1UJ3-HC 1YQV-HC 2D2H-HC 2CMR-HC 2DD8-HC 2QNHC 2QNHC 2QNL-HC 2QOL-HC 2QOL-HC 3D9A-HC 1BJ1-HC 1CE1-HC 1N82-HC	GQGTSVTVSSAKTTPPSVYPL GQGTSVTVSSAKTTPPSVYPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLITVSSASTTPSVYPL GQGTLITVSSASTTPSVYPL GQGTLITVSSASTTPSVYPL GQGTLITVSSASTTPSVYPL GQGTLITVSSASTTPSVYPL GQGTLVTVSSASTKGPSVYPL GQGTLVTVSSASTKGPSVYPL GQGTLVTVSSASTKGPSVYPL GQGTLITVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL GQGTLVTVSSASTKGPSVPPL	APGSJ APGSJ APGSJ APSSJ APSSJ APGCI APGSJ APGCI APGSJ APGSJ APGSJ APGSJ APGSJ APGSJ APSSJ APSSJ APSSJ	AAQTNSMVTLG GGGGQAWTIL SGTSGSTAAL KSTSGGTAAL GDTTGSSVTL ALKSNSMVTLG ALKSNSMVTLG GDTTGSSVTLG AAQTNSMVTLG GDTTGSSVTLG KSTSGGTAAL GDTTGSSVTLG KSTSGGTAAL GDTTGSSVTLG KSTSGGTAAL GDTTGSSVTLG KSTSGGTAAL KSTSGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTS KSTS KST	LVKGYFPEPVTV LVKGYFPEPVTV	"TWNSGSLSSGVHTFPA "TWNSGSLSSGVHTFPA "SWNSGALTSGVHTFPA "WNSGALSSGVHTFPA "TWNSGSLSSGVHTFPA "TWNSGSLSSGVHTFPA "TWNSGSLSSGVHTFPA "WNSGSLSSGVHTFPA "WNSGSLSSGVHTFPA "SWNSGALTSGVHTFPA "SWNSGALTSGVHTFPA "SWNSGALTSGVHTFPA "SWNSGLTSGVHTFPA "SWNSGALTSGVHTFPA "SWNSGALTSGVHTFPA "SWNSGALTSGVHTFPA "SWNSGALTSGVHTFPA "SWNSGALTSGVHTFPA "SWNSGALTSGVHTFPA "SWNSGALTSGVHTFPA "SWNSGALTSGVHTFPA "SWNSGALTSGVHTFPA "SWNSGALTSGVHTFPA "SWNSGALTSGVHTFPA "SWNSGALTSGVHTFPA "SWNSGALTSGVHTFPA "SWNSGALTSGVHTFPA "SWNSGALSGVHTFPA "SWNSGALSGVHTFPA "SWNSGALSGVHTFPA "SWNSGALSGVHTFPA "SWNSGALSGVHTFPA "SWNSGALSGVHTFPA "SWNSGALSGVHTFPA "SWNSGALSGVHTFPA "SWNSGALSGVHTFPA "SWNSGALSGVHTFPA "SWNSGALSGVHTFPA "SWNSGALSGVHTFPA "SWNSGALSGVHTFPA "SWNSGALSGVHTFPA "SWNSGALSGVHTPA "SWNSGALSGVHTFPA "SWNSGALSGVHTFPA "SWNSGALSGVHTFPA "SWNSGALSGVHTFPA "SWNSGALSGVHTPA "SWNSGN "SWNSGALSGVHTPA "SWNSGALSGVHTPA "SWNSGALSGVHTPA "SWNSGALSGVHTPA "SWNSGALSGVHTPA "SWNSGALSGVHTPA "SWNSGALSGVHTPA "SWNSGALSGVHTPA "SWNSGALSGVHTPA "SWNSGALSGVHTPA "SWNSGN "SWNSGN "SWNSGN "SWNSGN "SWN "SWN	/LQS-DLYTLSSSVTVPSS' /LQS-DLYTLSSSVTVPSS' /LQSSGLYSLSSVVTVPSS' /LQSSGLYSLSSVVTVPSS' /LQS-DLYTLSSSVTVPSS' /LQS-DLYTLSSVTVPSS' /LQS-GLYTLSSSVTVPSS' /LQS-GLYTLSSSVTVPSS' /LQS-GLYTLSSSVTVPSS' /LQS-GLYTLSSSVTVPSS' /LQS-GLYTLSSSVTVPSS' /LQSSGLYSLSSVVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS'	TWPSETVTONVAHE PWPSETVTONVAHE SLGTQTYIONVAHE SLGTQTYIONVAHE PRPSETVTONVAHE SLGTQTYIONVAHE SLGTKTYTONVAHE SLGTKTYTONVAHE TWPSQTVTONVAHE SLGTQTYIONVAHE	ASSTKVDKKIVPRDC ASSTKVDKKIVPRDC PASSTKVDKKVEPKSCDKT ASSTKVDKKIVPR ASSTKVDKKIVPR ASSTKVDKKIVPRC ASSTKVDKKIVPRCGGE ASSTKVDKKIVPRCGGE ASSTKVDKKIVPRCGGE ASSTKVDKKIVPRCGHE (PSNTKVDKKVEPKSCDKT ASSTKVDKKIAAA ASSTKVDKKIAAA ASSTKVDKKIAAA ASSTKVDKKIAAA	225 215 'HT 225 'HT 225 218 218 218 217 215 217 'SPLF 228 219 210 'HT 223 221 'HT 223 220 'HT 231 220 'HT	- 5 3 5 5 9 9 3 3 7 3 5 9 5 7 3 3 9 0 0 L 3 0 L 0 0 9
1FNS_HC 1HOD_HC 1IQD_HC 1JPS_HC 1KB5_HC 1KK3_HC 1UK3_HC 1UJ3_HC 1VQV_HC 1YQV_HC 1YQV_HC 2D27_HC 2D28_HC 2CMR_HC 2ON8_HC 2Q0N_HC 2Q0N_HC 2Q0N_HC 2Q0N_HC 2Q0N_HC 2Q1_HC 2QDR_HC 3D9A_HC 1BJ1_HC 1CE1_HC 1SY6_HC	GQGTSVTVSSAKTTPPSVYPL GQGTSVTVSSAKTTPPSVYPL GQGTMVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTKGPSVFPL GQGTLVTVSSASTTAPSVYPL GQGTLVTVSSASTTPPSVYPL GQGTLVTVSSASTKGPSVFPL GQGTLTVSSASTTPPSVYPL GQGTLVTVSSASTKGPSVFPL	APGSJ APGSJ APGSJ APSSJ APSSJ APGCI APGSJ APGCI APGSJ APGSJ APGSJ APGSJ APGSJ APGSJ APSSJ APSSJ APSSJ	AAQTNSMVTLG GGGGQAWTIL SGTSGSTAAL KSTSGGTAAL GDTTGSSVTL ALKSNSMVTLG ALKSNSMVTLG GDTTGSSVTLG AAQTNSMVTLG GDTTGSSVTLG KSTSGGTAAL GDTTGSSVTLG KSTSGGTAAL GDTTGSSVTLG KSTSGGTAAL GDTTGSSVTLG KSTSGGTAAL KSTSGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTSGGTAAL KSTS KSTS KST	LVKGYFPEPVTV LVKGYFPEPVTV	TWNSGSLSSGVHTFPA TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA 'SWNSGALTSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'TWNSGSLSSGVHTFPA 'SWNSGALTSGVHTFPA	/LQS-DLYTLSSSVTVPSS' /LQS-DLYTLSSSVTVPSS' /LQSSGLYSLSSVVTVPSS' /LQSSGLYSLSSVVTVPSS' /LQS-DLYTLSSSVTVPSS' /LQS-DLYTLSSVTVPSS' /LQS-GLYTLSSSVTVPSS' /LQS-GLYTLSSSVTVPSS' /LQS-GLYTLSSSVTVPSS' /LQS-GLYTLSSSVTVPSS' /LQS-GLYTLSSSVTVPSS' /LQSSGLYSLSSVVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS' /LQSSGLYSLSSVTVPSS'	TWPSETVTONVAHE PWPSETVTONVAHE SLGTQTYIONVAHE SLGTQTYIONVAHE PRPSETVTONVAHE SLGTQTYIONVAHE SLGTKTYTONVAHE SLGTKTYTONVAHE TWPSQTVTONVAHE SLGTQTYIONVAHE	ASSTKVDKKIVPRDC ASSTKVDKKIVPRDC ASSTKVDKKVEPKSCDKT ASSTKVDKKIVPR ASSTKVDKKIVPR ASSTKVDKKIVPRC ASSTKVDKKIVPRCGGU ASSTKVDKKIVPRCGGU ASSTKVDKKIVPRCGGU ASSTKVDKKIVPRCGHH (PSNTKVDKKVEPKSCDKT ASSTKVDKKIAAA ASSTKVDKKIAAA ASSTKVDKKIAAA ASSTKVDKKIAAA ASSTKVDKKIAAA ASSTKVDKKIAAA	225 215 'HT 225 'HT 225 218 218 218 217 215 217 'SPLF 228 219 210 'HT 223 221 'HT 223 220 'HT 231 220 'HT	- 5 3 5 5 9 9 3 3 7 3 5 9 5 7 3 3 9 0 0 L 3 0 L 0 0 9

Fig. 2 (continued).

# Propensity of Amino Acids to Occur in Binding Sites of Fabs and in APRs

Table III lists the antigen-binding residues of Fabs. The binding residues that also belong to APRs are highlighted in bold (Table III). The buried surface areas contributed by these binding residues are unequal. The binding residues are primarily located in CDR loops, with a few of them in the framework regions (FRs). Y49 in light chains is identified as a binding site residue in a few complexes. It immediately precedes L2 loops and, technically, falls in FR2. However, we pooled this residue with those in L2 in our analysis. Several residues in CDRs loops do not participate in antigen binding (Table IV). However, these

non-binding residues may help CDR loops acquire the proper backbone conformations essential for binding (49).

Fig. 5a and b compare the propensity of individual amino acid to occur in antigen-binding sites and in APRs in our dataset. Residues with propensity values above one are favored, while those with propensity values below one are disfavored. Our results are in general agreement with previous analysis on antibody-antigen interface (32,35). For example, Trp, Tyr and Arg, favored at antigen-binding sites in our analysis (Fig. 5a), are also the hot-spot residues for protein interfaces (50).

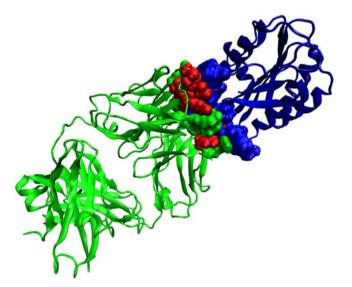
The APRs considered in propensity calculations are the APRs present in the overall sequences of Fabs including both variable (contributing towards binding) and constant

Number	PDB ID	Fab (Å <sup>2</sup> )	Antigen (Ų)	APRs in Fab <sup>a</sup> (%)	Number	PDB ID	Fab (Ų)	Antigen (Ų)	APRs in Fab <sup>a</sup> (%)
l	I FE8	970	1,012	29.8	16	2DD8	895	906	6.0
2	IFNS	628	703	3.7	17	2FD6	679	732	22.4
3	1H0D	699	731	2.4	18	2NXY	695	703	22.2
4	IIQD	1,064	1,265	10.8	19	2Q8B	I,303	1,291	35.8
5	IJPS	1,229	1,229	5.7	20	2QQN	468	464	42.7
6	IKB5	1,489	1,435	13.7	21	2R0L	1,078	1,128	12.3
7	ILK3	919	952	10.8	22	2VDR	997	1,020	15.7
8	IMLC	718	810	33.3	23	3D85	720	755	8.7
9	IOSP	845	756	0.0	24	3D9A	878	921	9.5
10	I UJ3	1,140	1,113	5.4	25	BJ	933	1,040	2.5
11	I WEJ	627	715	13.2	26	ICEI	369	557	16.6
12	IYQV	825	923	19.1	27	IN8Z	1,153	1,270	29.2
13	IZTX	755	808	20.2	28	ISY6	826	916	23.2
14	2B2X	I,480	1,500	17.2	29	IYY9	1,068	1,083	14.5
15	2CMR	1,188	1,186	18.5					

Table II Buried Surface Area of Fabs, APRs in Fabs and Antigens in Fab-Antigen Complexes

<sup>a</sup> Percent contribution towards buried surface area by the binding residues in Fab which also belong to APRs

domains (non-contributing ones) (Fig. 5b). Again our results are consistent with previous analyses of APRs (30,51,52). The aromatic residues (Tyr and Trp) and  $\beta$ -branched aliphatic residues Val, Ile, and Leu are favored in APRs. Charged residues are strictly avoided in APRs. Taken together, aromatic residues, Tyr and Trp, are favored both in antigen-binding sites and APRs. Hence, these residues may be the coupling agents between aggregation and antigen recognition.



**Fig. 3** Ribbon representation for the structure of IGG RU5 Fab (green)-Von Willebrand factor (blue) complex (PDB ID: IFE8). Only the binding residues in the Fab-antigen complex identified by our method are shown in CPK representation. Binding residues in Fab that also belong to APRs are shown in red color.

# Buried Surface Area Contributions and Coincidence of APRs with Individual CDR Loops

Fig. 6a shows the contribution of each CDR loop towards the buried surface area of the Fabs in the complexes. The average values along with the number of binding residues and CDR length are also presented in Table IV. The buried surface area values are very similar to those obtained by Wilson and Stanfield (1994) (53). On average, H3 loops contribute the most towards antigen binding  $(25\pm13.8\%)$ , Table IV). The portion of residues that directly contact antigens is also the greatest (48%) for H3 loops. If the average buried surface area contribution for each loop is normalized by number of binding residues in the loop, H3 loops again have the greatest contribution (5.3% per residue). Consistently, the variability plots of antibody sequences and statistical studies of antigen-contacting residues have shown that H3 loops are important for antibody specificity (36, 54, 55).

The average contribution of H2 loops is almost equal to that of the H3 loops  $(23.5\pm10.9\%;$  Table IV). H2 loops are the longest among CDR loops  $(16.9\pm0.6 \text{ residues})$ , and they contribute the most residues  $(5.7\pm2.3)$  towards antigen binding (Table IV). The average buried surface area contribution by L3 loops is the third largest  $(17.3\pm8.6\%)$ . Taken together, these three loops, namely, H3, H2 and L3, contribute approximately two-thirds  $(65.8\pm17.7\%, \text{ range:}$ 19-100% of the buried surface area of the Fabs. The shortest CDR loops, L2, also contribute the least  $(7.4\pm$ 7.5%, Table IV). In fact, these loops do not contribute towards the buried surface area of Fabs in seven (24%)complexes in our dataset (Fig. 6 and Table III). The overall

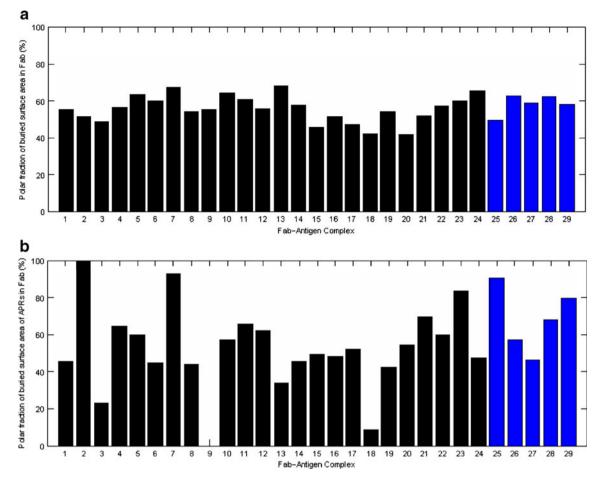


Fig. 4 (a) Polar fraction of the buried surface area of Fabs for all complexes in our dataset. (b). Polar fraction of the buried surface area of APRs in Fab. The last five commercial Fab-antigen complexes are in blue color.

contribution of the CDR loops in heavy chain (CDR loops: H1+H2+H3) towards buried surface area of Fab is  $60.7 \pm 11.3\%$  (range 43–86%). Only in 6 out of 29 (approximately 21%) complexes, the heavy chain CDR loops contribute less than 50%. We have also counted the numbers of hydrogen bonds and ion pairs between Fab and antigen to study the specific interactions at Fab-antigen interfaces (Table IV). On average, the trend is the same. H2, H3 and L3 loops again make the greatest number of hydrogen bond and ion pair contacts with the antigens.

Fig. 6(b) plots the frequency of coincidence of APRs and binding residues in CDR loops. These values estimate the potential coupling between aggregation and antigen recognition for each CDR loop. Residues in APRs are also among the antigen-binding residues in CDR loops with an average frequency of  $29.3 \pm 12.5\%$ . The two most important antigenbinding contributors, H2 and H3, show very different aggregation coupling frequencies. H2 loop has the highest APR incidence (44.8%), while H3 loop shows the lowest incidence (6.9%). The coincidences of APRs and binding residues in other CDR loops, L1, L2, L3 and H1, are similar ( $31.0 \pm 2.8\%$ ).

#### DISCUSSION AND CONCLUSION

In the literature, there is increasing consensus that aggregation is an intrinsic property of proteins. In our previous work, we found that commercially available therapeutic mAbs contain short APRs similar to those seen in the amyloidogenic proteins (30). Interestingly, some of the potential APRs in mAbs overlapped with the CDR regions. However, that study could not tell if the APRs localized in the CDRs in those mAbs actually contributed to antigen recognition also. Present survey shows that CDR-localized potential APRs contribute significantly towards antigen recognition. This suggests the possibility of linkage between aggregation and loss of function in antibody-based therapeutics. This link is undesirable from the pharmaceutical point of view because therapeutic mAb drug substances as well as products are stored without their cognate antigens over long periods of time.

The link between aggregation and CDR loops has been previously reported based on experimental studies of domain antibodies (56). Moreover, APRs have been shown to overlap with protein-protein interfaces in general (57).

	Number	PDB ID			[]		El		Ī			H2			H3					Non-CDR	
	_			530 '32			G91 N94			10 T3I Y			D54	N56		797				S67 (LC)	
100         100         100         100         100         100         100         100         100         100         100           101         101         101         101         101         101         101         101         101         101         101         101           101	2			'32			E92		96/						A97 [			G100			
M M M M M M M M M M M M M M M M M M M				ç			0			c					N 100A		Y100D				
100         33         74         60         13         74         73         60         73         60         73         60         73         60         73         60         73         60         73         60         73         60         73         60         73         60         73         60         73				728			591 V94			m							Y100B				
1         1				30	Y49		G92			Ē	Ņ						797			ō	V2 (HC)
10121				'32	R54				РЗ				~				101C				
10         12         12         13<	5			328	Y50	E56	16H						E53	N56				Y100]			
163         173         174         173         174         175 <td></td> <td></td> <td></td> <td>'32</td> <td></td> <td></td> <td>S94</td> <td></td> <td>-</td> <td></td> <td></td> <td></td> <td>Υ59</td> <td>264</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td>				'32			S94		-				Υ59	264							
61         13         15<	9	IKB5		130	Y49		16H						G55	157			96C			DI (LC)	
$ \left[ \begin{array}{cccccccccccccccccccccccccccccccccccc$				r32	T53	E56	Т94	У96	G				06				Y 100J				
T30R31T30R31K31K36K3				:28	K50		16S			3			N53	Y56			398			DI (LC)	
$ \left[ \begin{array}{cccccccccccccccccccccccccccccccccccc$				<b>1</b> 32			G94		96				Y59			VI 00					
105         131         132         133 <td></td> <td></td> <td>N32</td> <td></td> <td>Y50</td> <td></td> <td>16S</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>G53</td> <td></td> <td></td> <td></td> <td>398</td> <td></td> <td></td> <td></td> <td></td>			N32		Y50		16S						G53				398				
105         73         81         82         63         79         793         734         746         733         733         734         745         733         733         734         743         733         733         734         743         733         733         734         733         734         733         734         733         734         733         734         733         734         733         734         733         734         733         734         734         733         734         733         734         733         734							W94	R96	ž												
	6						16Y	W92	S3	_		Υ53							0015	D66 (LC)	
Iwe         Had         Nat         S94         Y94         Y34         Y35         Y35 <td>01</td> <td></td> <td></td> <td></td> <td></td> <td>T52</td> <td>16H</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>N54</td> <td></td> <td></td> <td></td> <td></td> <td>Y98</td> <td></td> <td>DI (LC)</td> <td></td>	01					T52	16H						N54					Y98		DI (LC)	
IVM         H3         Y3         K3							S94	У96	ŝ				Y59				398				
IrQu         N3I         Y32         D50         W31         Y32         D50         W31         S54         G55         S4         G55         S4         G55         S4         S54         S4         S54         S45         S44         S47         S44         S47         S44         S47         S44         S45         S45         S45         S45         S45         S45         S45         S45         S47         S44         S47         S44         S47         S44         S45         S44         S44         S45         S44         S44         S45         S44         S44         S45         S45	=			431 Y32		K52	W92		Q	D31	33		S54	N56							
IZTX       530       V30       V33       V31       V32       V33       V31       V31       V32       V31       V31       V31       V32       V31       V31       V31       V31       V32       V31       V31       V32       V31	12			r32	D50		16M			D3			S54	355							
IZTX         50         WG0         H91         Y92         Y27         T8         D31         L5         R94         S95         S97           2B2X         Q29         V30         Y49         L5         S32         Y33         W33         Y32         W33         Y32         W33         Y33         Y32         W33         Y33         Y34         Y36         Y37         Y33         Y34         Y36         Y33         Y34         Y33         Y34         Y33         Y34         Y33         Y34         Y33         Y34         Y36         Y37         Y33         Y34         Y33         Y34         Y34         Y33         Y34         Y34         Y33         Y34																		D98		W47 (HC)	
101         103 <td>13</td> <td></td> <td>S30</td> <td></td> <td>W50</td> <td></td> <td>16H</td> <td>Y92</td> <td>2</td> <td>T28</td> <td>31</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>597</td> <td></td> <td></td> <td></td> <td></td>	13		S30		W50		16H	Y92	2	T28	31						597				
2B2X         Q2         V30         V40         L50         S2         T28         S30         R31         V50         S52         G33         R94         F96         G97         V2           N31         H32         L54         A5         S56         N94         Y96         732         T33         G54         H56         Y88         D98         Y100         D101           2CMR         V30         H31         K50         Y91         S92         S30         S31         T52         T53         T54         H56         T58         N96         P70         D101           M32         M31         K50         S31         S31         T52         T53         T54         T58         T58         T59         P70         P71         P73         H71           M32         M31         Y32         M33         T52         T53         T54         T56         T58         T56         V97         V91         F73         H71           M32         Y33         Y32         Y33         T52         T53         T56         V97         V97         V93         Y94         Y93         Y94         Y93         Y94         Y93							T93	Т94	£								0010				
N31         H32         L54         A55         S56         N94         W96         Y32         T33         G54         H56         Y58         D98         Y100         D101           2CMR         Y30         H31         K50         Y91         S92         S30         S31         I52         I53         F54         G55         D95         N96         P77           W32         W33         Y34         Y32         X33         T56         A57         N58         Y59         T96         P71         M98           2DD8         N27         S30         X31         T32         T33         T52         I53         L54         D95         T96         V71         M98           X91         X91         Y32         T33         T52         I53         L54         D95         T96         V71         M98           X95         Y96         Y96         Y96         Y32         T33         T52         I53         L54         D95         T96         V71         M98				/30	Y49			S92	Т2	S30							297			V2	<b>W47</b> (HC)
2CMR         Y3         H3         K50         Y9         S92         S30         S31         I52         I53         F54         G55         D95         N96         P97           W32         W33         Y94         Y32         A33         T56         A57         N58         Y59         T98         L100           2DD8         N27         S30         Y31         Y32         T33         T52         I53         L54         D95         T96         V97         M98           D95A         Y96         Y32         T33         T52         I53         L54         D95         T96         V97         M98				-132	L54		N94	96M	£								1010				
W32         N93         Y94         Y32         A33         T56         A57         N58         Y59         T98         L100           2DD8         N27         S30         W91         S93         S94         S31         Y32         T33         T52         I53         L54         D95         T96         V97           D95A         Y96         T33         T52         I53         L54         D95         T96         V97				131	K50		16Y	S92	S				F54				26c			E73 (HC)	
2DD8 N27 S30 W91 S93 S94 S31 <b>Y32 T33</b> T52 I53 L54 D95 T96 V97 D95A Y96 I56 N58			W32				N93	Ү94	¥				N58	r59		001					
Y96 I56				30			16M			Y32	33							M98			
							D95A						58								

	Q3 L4 S74									
	V2 N76 (HC) G68 /1 ()							V2(HC)		(TC)
								R66(LC)		
	EI S75 S67	00				G100		R6(		
								8		
Q	8		0			66Y	0	7 D98 2		
EI 00D	F102		УI 00	V98		798 Г	A100			
W99 G100C N100G		96M	66人	D97		H97 W/100D	H97	G96 Y100A	У100	799
H98 E100B D100F	<b>596</b>	M100B F95	R98 D98	96M	D97	P96 SLOOR	506 G96	W95 F100	D98	Y98 Y100A
W95 E99 Y100E		<	797 797	N95	96M	Υ95 S100A			D97	T97 D100
> ü >		G55 R F	> >	2	>	> v	LL C	₩ 0		
N56			K64							
D55 Ү59 <b>L54</b>	A53	T53	D58 N54	Y53 K58	S54	T54				G55
E52 E58 E58 H58	D52		T57 D52 K58		Y53 Ү58		K53		S53 N58	G54 D58
W50 T57 I52 V56	R50	W50	A56 R50 Y56	<b>Y50</b> N54	Υ50 S56	~			N52 Y56	2 S53 T57
	K28	10				W50		I R50 T57	Υ50 R54	<b>W52</b> N56
m	<b>6 F27</b> 2 S33	H35	5 Y33	5 N32	8 8	53		D3	0	_
31 X33	<b>5 G26</b>	31 S32	81 T32	0 S31 3 H35				2 Y33	8 T30 - T33	D N N N
N31	<b>S25</b> D31	T93 N3I	D31	S93 T30 V33	S93 D27 D32		593 Y33	G26 T94 T32	T28 R31	<b>N93</b> T30
N92 F95 P95	R96		P96 A92 L94			592 M96	192 St R96		S92 N94	_
W91 N Y93 F W94 F	G91 F		791 F 791 <i>j</i> 093 l						W91 S	
/ / /	•	,		•	01 >	• •		-	> 01	- /
	T56		•					F53		
	<b>Y50</b>		) Y50 3 D56		3 Y50		N50		0	-
K53	<b>Y49</b> R54		Y49 N53	Υ50	K49 O53	,	Y49	Y49 Y55	D50	Y50
Y32	V34	<b>732</b> S30	N32		N3I			<b>A32</b>	Y32	S28
<u>3</u> 31	D32	531 D28	<b>T3</b>	Y32	G30 N32	(EX	Y32	D28 T3I	23	Q27 N32
2FD6 2NXY	2Q8B 200N	2ROL	2VDR	3D85	3D9A	IBI	ICE I	Z8NI	ISY6	6771
	61 00	21 20	22	23	24	25	26	27	28	29

CDR Statistics	LI	L2	L3	ΗΙ	H2	H3	FR
Length of CDR loop (L)	12.0±0.9	$7.0\pm0.0$	$9.0\pm\!0.5$	$10.0 \pm 0.0$	16.9±0.6	$9.8 \pm 3.0$	
Number of binding residues ( $N_b$ )	$2.6 \pm 1.2$	$2.0\pm2.1$	$4.0 \pm 1.4$	3.1±1.8	$5.7\pm2.3$	$4.7 \pm 1.8$	$0.8\pm1.4$
Ratio I (%, N <sub>b</sub> / L)	21.7	28.6	44.4	31.0	33.7	48.0	
Percentage of buried surface area of Fab (SA buried)	$12.3\pm8.4$	$7.4\pm7.5$	$17.3 \pm 8.6$	2. ±8.3	$23.5\pm10.9$	$25.0\pm13.8$	$2.3\pm5.1$
Ratio 2 (SA buried / N <sub>b</sub> )	4.7	3.7	4.3	3.9	4.1	5.3	
Number of hydrogen bonds and ion pairs	$1.6 \pm 1.6$	$0.7 \pm 1.0$	2.1±1.7	$1.8 \pm 1.7$	$3.5\pm2.6$	3.1±2.0	

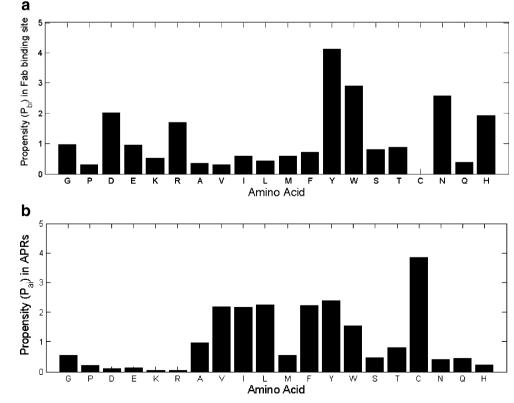
Table IV Contribution of Individual CDR Loop Towards Antigen Recognition

Here, our survey seeks to understand the detailed characteristics of this link in the context of antibody-antigen recognition. Interestingly, we also observed that the APRs in Fc region of the mAbs lie close by but do not overlap with various protein (*viz*.  $F_c\gamma$  receptor(s), FcRn, Protein A and Protein G) binding sites. For example, APRs 269-VTCVV-274, 285-FNWYV-289 and 312-VVSVLTVL-319 lie close to but do not overlap with  $F_c\gamma$ RIII and C1q binding sites in the crystal structure of the human antibody IgG1 b12 against HIV-1 (PDB code: 1HZH) (58).

Computational approaches towards identification of potential APRs utilize sequence (19,28,29,59,60) and structural (21) methods. In numerous instances, these predictions have been experimentally validated and used to design peptides and proteins with lower aggregation propensities (61,62). The sequence-based approaches have similar levels of accuracy as 3D profile-based ones (16). The

sequence-based approaches are faster and require less computational resources as compared to the structural ones. These can be very useful in early discovery stage, where a large number of sequences are screened for potential leads. TANGO and PAGE are two such computational programs that require only the protein sequence as input. The prediction from TANGO relies on physicochemical rules behind  $\beta$ -sheet formation and assumes that the core of the aggregates is completely desolvated (28). It also takes into account the competition between other conformational states, such as  $\alpha$ -helix,  $\beta$ strand, turn, random coil and  $\beta$ -aggregates for the given sequence region. PAGE computes the aggregation propensity based on aromaticity,  $\beta$ -strand propensity, charge, solubility and hydrophobicity of the residues in window of 5-9 residues along the sequence (29). Both TANGO and PAGE have been used on a large number of peptides and

Fig. 5 (a) Propensity  $(P_{bi})$  of individual amino acid to occur at binding sites in Fab. (b) Propensity  $(P_{ai})$  of individual amino acid to occur in APRs in Fab.



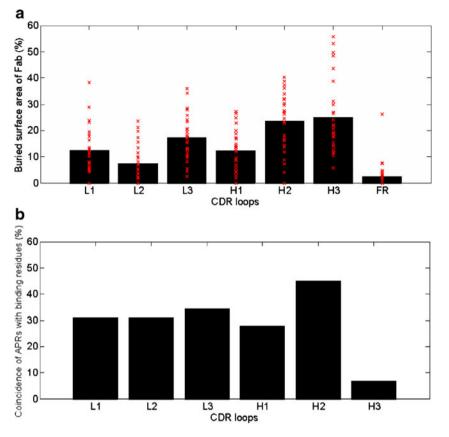


Fig. 6 (a) Contribution of each CDR loop towards the buried surface area of Fab. Contributions from all framework regions (FR) are pooled and are shown as FR contribution. The bars show the average values over all complexes. The red crosses show the individual values for the 29 complexes. (b) Coincidence between APRs and binding residues in CDR loops. If a CDR contains at least one binding residue which also belongs to an APR, it is counted as one incidence. For example, the coincidence between APRs and binding residues of the 29 complexes (31%) contain at least one binding residue which also belongs to APR.

small proteins and show high prediction accuracies. TANGO was found be more than 92% accurate in experimental validation studies (28). TANGO has also been used extensively by other groups (62–66). We used more than one program to identify APRs so that our results are not influenced by peculiarities of training sets and algorithms used. We noted good overlap between APRs of type 1 and type 2, indicating the general agreement between the two programs.

Molecular simulation techniques have also been used to study aggregation (22,23,25,26,67-69). This approach was recently applied to full-length antibody by Chennamsetty *et al.* (2009) (23,24). They have developed a technique named *spatial-aggregation-propensity* (SAP). SAP uses residue hydrophobicity and dynamical conformations collected from simulations to identify surface exposed hydrophobic patches. This concept is similar to the surface aggregation propensity used by Pechmann *et al.* (2009) (57). These surface-exposed hydrophobic patches can act as structural hot-spots for aggregation. These patches look for non-polar residues in close spatial vicinity on protein surfaces. The spatially vicinal residues do not have to be contiguous in amino acid sequence. Moreover, these surface patches are not necessarily potential cross  $\beta$ -aggregation sites. Hence, they are different from the cross  $\beta$ -motif-forming APRs identified by the sequence-based approaches. However, APRs also often contain hydrophobic  $\beta$ -branched aliphatic and aromatic residues. Hence, the APRs may overlap with the surface hydrophobic patches but not be completely equivalent to them. APRs can also be significantly polar, especially those containing Asn and Gln residues. The examples include yeast prion proteins sup35 and Ure2p and several neurodegenerative diseases, like Huntington's disease (15,70–72). In the Fabs of our dataset, the APRs in L3 loops are polar (Fig. 2a). Overall, the APR and SAP approaches are complementary to each other.

To promote aggregation, an APR should have high intrinsic aggregation propensity, be conformationally unstable/flexible, be surface exposed or become exposed upon conformational transition and facilitate intermolecular interactions. Hence, three-dimensional structures are important for identifying which of the potential APRs could really initiate self-association. In a recent study, Hamada *et al.* (2009) studied the ability of individual  $\beta$ -strands to initiate amyloid-like fibril formation in  $\beta$ -lactoglobulin (73). They found that sequence regions with high intrinsic aggregation propensity still need at least local unfolding in the native structure to be able to seed aggregation. In light of this, one could postulate that the potential APRs in these surface-exposed CDR loops could seed aggregation in the therapeutic mAbs *via* self-association of Fabs. This process may be materially assisted by two factors: first, the mobility of the CDR loops is enhanced due to the absence of cognate antigens; second, the physico-chemical stresses may further perturb the native Fab structures locally.

Besides the strong experimental evidence for the existence of short sequentially contiguous cross  $\beta$ -aggregate-forming regions in general proteins, the available experimental evidence suggests that APRs may also play a similar role in biotherapeutic aggregation as well. Some of the evidence is discussed below.

*I*. Biopharmaceuticals at the end of expiration date were reported to form aggregates which bind Thioflavin T and Congo Red (27). Thioflavin T-binding is the characteristic of aggregates containing cross-β structures (7). Our previous analysis showed that the biopharmaceuticals reported by Maas *et al.* (2007) contain several TANGO/PAGE predicted APRs (See Table III in Reference 30) (30). Moreover, the TANGO/PAGE predicted APR 14-ALYLV-18 coincides with the experimentally proven fibril-forming segment 12-VEALYL-17 of insulin (21). Similar results were observed by us upon exposure of several different IgG2 mAbs to thermal stress (internal effort; data not shown).

2. Several experimental reports that study aggregation in the proteins and document the sequence changes in their variants (homologues or mutants) with reduced aggregation propensities were found to disrupt/mitigate the APRs predicted by TANGO and PAGE in our analyses. In particular, we cite three examples because of good agreement between the experiments and computational predictions. These are bovine growth hormone (74), amyloidogenic immunoglobulin light chain (75) and a human IgG1 mAb (23,24). We summarize our findings below; the details of these cases are presented in Supplementary Material. First, based on the experimental studies, Lehrman et al. have identified a sequence region 109-133 in bovine growth hormone (bGH) to be involved in aggregation (74). Our TANGO and PAGE analyses indicate the presence of APRs in this region (119-GILALM-124). The experimentally designed variants 8H-bGH and human growth hormone (hGH) differ in the sequence region 109-133 and show reduced aggregation propensity in the experiments of Lehrman et al. (74). Consistent with this observation, TANGO/PAGE spectra do not indicate an APR in this region for 8H-bGh and hGH. Second, Baden et al. (2008) have identified three non-conserved somatic mutations, I34N, Q42K and H87Y, which restore amyloidogenic immunoglobulin light chain AL-09 to its germline sequence KIO18/O8 (75). TANGO spectrum of AL-09 indicates a very strong APR containing I34 (32-YLIWY-36, TANGO aggregation score ~90%). The PAGE spectrum shows an APR containing H87 (87-HCOOY-91). Both TANGO and PAGE did not find an APR involving the residues at the position 42. The corresponding TANGO spectra for both the germline sequence  $\kappa$ IO18/O8 and the AL-09 mutants with reduced aggregation propensity indicate that APR 32-YLIWY-36 is substantially weakened for the germ line light chain due to the somatic disruption at position 34. However, APR 87-HCOOY-91 identified by PAGE remained unchanged. Third, L309K mutation in the C<sub>H</sub>2 domain of intact IgG1 mAb studied by Chennamsetty et al. (2009) reduces its aggregation propensity as shown by turbidity and HPLC assays and improves its stability in DSC experiments (see Table I and Fig. 3 in Ref. 24) (24). This mutation actually disrupts a strong APR (302-VVSVLTVL-309) (TANGO aggregation score ~90%) found in the Fc regions of IgG mAbs. This APR is well conserved among immunoglobulin Gs and was documented earlier (30). In all three cases, both TANGO and PAGE also detected additional APRs that were common between the proteins and their variants, indicating the potential for further reduction in aggregation propensities of the molecules.

**3**. Use of TANGO/PAGE predictions in combination with the molecular modeling can help identify positions/mutations to reduce aggregation propensities in the biotherapeutics. Recently, we observed that disruption of a TANGO-predicted strong APR in FR2-L2 region of an IgG2 mAb by a single point mutation reduced its aggregation propensity and improved solubility as indicated by biophysical experiments (in-house effort; data not shown).

Similar to small proteins and peptides, the potential APRs in the antibodies are also short sequence regions. Hence, one or a few changes in sequences that disrupt the CDR-localized APRs may significantly reduce aggregation propensity of the mAbs. From a product formulation and developability point of view, improvement in mAb stability and solubility is desirable. This may help improve expression levels in cell lines, facilitate high concentration dosage forms and increase shelf-life of the product. However, developability-related sequence mutations should not adversely impact the potency of the therapeutic mAbs. In this regard, the observed incidence of potential APRs in the CDR loops and adjacent framework β-strands is significant. Disruption of these APRs without affecting therapeutic mAb potency could be difficult and time-consuming without a rational approach. A structure-based input that simultaneously considers all these issues may lead to more "druggable" therapeutic candidates. The present study offers useful guidelines for drug candidate design and selection at early discovery and formulation stages:

1. Aromatic residues Tyr and Trp are favored both in APRs and CDRs (Fig. 5). Aromatic amino acids have been known to play an important role in directing molecular recognition, mostly because of their ability to form  $\pi$ -stacking interactions (76). Tyr is also frequently used in high affinity protein-protein interface design (77). Bogan *et al.* have reported that hot-spots of binding energy at protein interfaces are rich in Trp and Tyr (50). Hence, caution should be used when considering mutation of Tyr and Trp residues in mAb CDRs to alleviate aggregation because it could also impact mAb potency. Instead, mutation of a sequence neighbor to disrupt the APR's amyloidogenic sequence pattern may be more appropriate.

2. The disruption of APRs should be performed in such a way that the CDR loop conformations, especially the conformations of residues that contribute substantially to antigen recognition, are not disturbed. Molecular models, crystal structures of Fab-antigen complexes and Alanine scanning experiments may help identify such binding hotspot residues. Mutations affecting these residues should be avoided to preserve potency. On average, CDR loops H3, H2 and L3 contribute the most towards antigen recognition. H3 loops do contain fewer APRs, but this is not the case for H2 and L3 loops. While different binding site residues contribute unequally towards antigen recognition, there is still some risk that mutations in these loops could also impact potency of the mAb candidates. This is especially true at early discovery stages where Fab-antigen complex crystal structures are not yet available, antigen-binding residues are not well identified, and project time-lines are tight. On the other hand, L2 loops contribute the least towards antigen binding but often contain APRs (Fig. 6). Mutations aimed at disrupting the APRs in this region could help improve mAb stability without significantly impacting potency. Apart from L2 loops, L1 and H1 loops are the other regions where disruption of potential APRs can reduce mAb aggregation tendency and not impact the potency substantially.

**3.** Hydrophobic residues, especially Val, Ile and Leu, have high propensities for aggregation but not for antigen binding (Fig. 5). Hence, APR disruption *via* mutation of these residues to polar or charged residues should be considered, especially when these residues are close to the CDR regions and are surface exposed.

**4**. APR disruption in the constant regions of the mAbs can also reduce the aggregation propensity. However, in case of therapeutic mAbs, caution is advised because of the potential for drifting from germ-line sequences.

#### ACKNOWLEDGEMENTS

We thank Drs. Sandeep Nema, Sa V Ho, Graeme Bainbridge, Tapan Das, Gerald Casperson and Satwik Kamtekar for several helpful discussions. A postdoctoral fellowship for Xiaoling Wang at BioTherapeutics Pharmaceutical Sciences, Pfizer, is gratefully acknowledged.

#### REFERENCES

- 1. Woodnutt G, Violand B, North M. Advances in protein therapeutics. Curr Opin Drug Discov Devel. 2008;11:754–61.
- Hudson PJ, Souriau C. Engineered antibodies. Nat Med. 2003;9:129–34.
- Foulkes R. Preclinical safety evaluation of monoclonal antibodies. Toxicology. 2002;174:21–6.
- Cromwell MEM, Hilario E, Jacobson F. Protein aggregation and bioprocessing. AAPS J. 2006;8:E572–9.
- Manning MC, Patel K, Borchardt RT. Stability of protein pharmaceuticals. Pharm Res. 1989;6:903–18.
- Rosenberg AS. Effects of protein aggregates: an immunologic perspective. AAPS J. 2006;8:E501–7.
- Gebbink MF, Bouma B. Method for detecting and/or removing protein comprising a cross-beta structure from a pharmaceutical composition. U.S.P.T. Office US20070015206A1 (2007).
- de Groot N, Pallares I, Aviles F, Vendrell J, Ventura S. Prediction of "hot spots" of aggregation in disease-linked polypeptides. BMC Struct Biol. 2005;5:18.
- Chiti F, Taddei N, Baroni F, Capanni C, Stefani M, Ramponi G et al. Kinetic partitioning of protein folding and aggregation. Nat Struct Mol Biol. 2002;9:137–43.
- Ventura S, Zurdo J, Narayanan S, Parreño M, Mangues R, Reif B *et al.* Short amino acid stretches can mediate amyloid formation in globular proteins: The Src homology 3 (SH3) case. Proc Natl Acad Sci USA. 2004;101:7258–63.
- Ivanova MI, Sawaya MR, Gingery M, Attinger A, Eisenberg D. An amyloid-forming segment of β2-microglobulin suggests a molecular model for the fibril. Proc Natl Acad Sci USA. 2004;101:10584–9.
- Monsellier E, Ramazzotti M, de Laureto PP, Tartaglia G-G, Taddei N, Fontana A *et al.* The distribution of residues in a polypeptide sequence is a determinant of aggregation optimized by evolution. Biophys J. 2007;93:4382–91.
- Murphy RM. Peptide aggregation in neurodegerative disease. Annu Rev Biomed Eng. 2002;4:155–74.
- Nelson R, Sawaya MR, Balbirnie M, Madsen AO, Riekel C, Grothe R *et al.* Structure of the cross-β spine of amyloid-like fibrils. Nature. 2005;435:773–8.
- Nelson R, Eisenberg D. Recent atomic models of amyloid fibril structure. Curr Opin Struct Biol. 2006;16:260–5.
- Kumar S, Wang X, Singh SK. Identification and impact of aggregation prone regions in proteins and therapeutic mAbs. In: Wangand W, Roberts C, editors. Aggregation of therapeutic proteins. US: Wiley; 2010. In press.
- Wang L, Maji SK, Sawaya MR, Eisenberg D, Riek R. Bacterial inclusion bodies contain amyloid-like structure. PLoS Biol. 2008;6:1791–801.
- Chiti F, Stefani M, Taddei N, Ramponi G, Dobson CM. Rationalization of the effects of mutations on peptide and protein aggregation rates. Nature. 2003;424:805–8.
- Tartaglia GG, Pawar AP, Campioni S, Dobson CM, Chiti F, Vendruscolo M. Prediction of aggregation-prone regions in structured proteins. J Mol Biol. 2008;380:425–36.

- Zhang Z, Chen H, Lai L. Identification of amyloid fibril-forming segments based on structure and residue-based statistical potential. Bioinformatics. 2007;23:2218–25.
- Thompson MJ, Sievers SA, Karanicolas J, Ivanova MI, Baker D, Eisenberg D. The 3D profile method for identifying fibril-forming segments of proteins. Proc Natl Acad Sci USA. 2006;103:4074–8.
- Cecchini M, Curcio R, Pappalardo M, Melki R, Caflisch A. A molecular dynamics approach to the structural characterization of amyloid aggregation. J Mol Biol. 2006;357:1306–21.
- Chennamsetty N, Helk B, Voynov V, Kayser V, Trout BL. Aggregation-prone motifs in human immunoglobulin G. J Mol Biol. 2009;391:404–13.
- Chennamsetty N, Voynov V, Kayser V, Helk B, Trout BL. Design of therapeutic proteins with enhanced stability. Proc Natl Acad Sci USA. 2009;106:11937–42.
- Vitalis A, Wang X, Pappu RV. Atomistic simulations of the effects of polyglutamine chain length and solvent quality on conformational equilibria and spontaneous homodimerization. J Mol Biol. 2008;384:279–97.
- Ma B, Nussinov R. Simulations as analytical tools to understand protein aggregation and predict amyloid conformation. Curr Opin Chem Biol. 2006;10:445–52.
- Maas C, Hermeling S, Bouma B, Jiskoot W, Gebbink MFBG. A role for protein misfolding in immunogenicity of biopharmaceuticals. J Biol Chem. 2007;282:2229–36.
- Fernandez-Escamilla A-M, Rousseau F, Schymkowitz J, Serrano L. Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. Nat Biotechnol. 2004;22:1302–6.
- Tartaglia GG, Cavalli A, Pellarin R, Caflisch A. Prediction of aggregation rate and aggregation-prone segments in polypeptide sequences. Protein Sci. 2005;14:2723–34.
- Wang X, Das TK, Singh SK, Kumar S. Potential aggregation prone regions in biotherapeutics: A survey of commercial monoclonal antibodies. mAbs. 2009;1:1–14.
- Berman H, Henrick K, Nakamura H. Announcing the worldwide Protein Data Bank. Nat Struct Mol Biol. 2003;10:980.
- Jackson RM. Comparison of protein-protein interactions in serine protease-inhibitor and antibody-antigen complexes: Implications for the protein docking problem. Protein Sci. 1999;8:603–13.
- Kortemme T, Baker D. A simple physical model for binding energy hot spots in protein-protein complexes. Proc Natl Acad Sci USA. 2002;99:14116–21.
- Tsai C-J, Lin SL, Wolfson HJ, Nussinov R. A dataset of proteinprotein interfaces generated with a sequence-order-independent comparison technique. J Mol Biol. 1996;260:604–20.
- Lo Conte L, Chothia C, Janin J. The atomic structure of proteinprotein recognition sites. J Mol Biol. 1999;285:2177–98.
- MacCallum RM, Martin ACR, Thornton JM. Antibody-antigen interactions: contact analysis and binding site topography. J Mol Biol. 1996;262:732–45.
- Jones S, Thornton JM. Principles of protein-protein interactions. Proc Natl Acad Sci USA. 1996;93:13–20.
- Janeway CA, Travers P, Walport M, Shlomchik MJ (Eds.). Immunobiology: The immune system in health and disease. 6th edition. p 683, Garland Science: New York, USA, 2004.
- Kabat EA, Wu TT, Perry HM, Gottesman KS, Foeller C. Sequences of proteins of immunological interest. Public Health Service, National Institues of Health, 1991.
- Labute P. The generalized Born/volume integral implicit solvent model: estimation of the free energy of hydration using London dispersion instead of atomic surface area. J Comput Chem. 2008;29:1693–8.
- 41. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM *et al.* A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. J Am Chem Soc. 1995;117:5179–97.

- Lee B, Richards FM. The interpretation of protein structures: estimation of static accessibility. J Mol Biol. 1971;55:379–80.
- Sticke DF, Presta LG, Dill KA, Rose GD. Hydrogen bonding in globular proteins. J Mol Biol. 1992;226:1143–59.
- Kumar S, Bansal M. Dissecting α-helices: position-specific analysis of α-helices in globular proteins. Proteins. 1998;31:460–76.
- 45. Braden BC, William Dall'Acqua AC, Fields BA, Goldbaum FA, Malchiodi EL, Mariuzza RA *et al.* Structure and thermodynamics of antigen recognition by antibodies. Ann NY Acad Sci. 1995;764:315–27.
- Chothia C, Janin J. Principles of protein-protein recognition. Nature. 1975;256:705–8.
- Lippow SM, Wittrup KD, Tidor B. Computational design of antibody-affinity improvement beyond *in vivo* maturation. Nat Biotechnol. 2007;25:1171–6.
- Sinha N, Mohan S, Lipschultz CA, Smith-Gill SJ. Differences in electrostatic properties at antibody antigen binding sites: implications for specificity and cross-reactivity. Biophys J. 2002;83:2946–68.
- Chothia C, Lesk AM, Tramontano A, Levitt M, Smith-Gill SJ, Air G *et al.* Conformations of immunoglobulin hypervariable regions. Nature. 1989;342:877–83.
- Bogan AA, Thorn KS. Anatomy of hot spots in protein interfaces. J Mol Biol. 1998;280:1–9.
- Bemporad F, Taddei N, Stefani M, Chiti F. Assessing the role of aromatic residues in the amyloid aggregation of human muscle acylphosphatase. Protein Sci. 2006;15:862–70.
- 52. Kim W, Hecht MH. Generic hydrophobic residues are sufficient to promote aggregation of the Alzheimer's Aβ42 peptide. Proc Natl Acad Sci USA. 2006;103:15824–9.
- Wilson IA, Stanfield RL. Antibody-antigen interactions: new structures and new conformational changes. Curr Opin Struct Biol. 1994;4:857–67.
- 54. Kabat EA, Wu TT. V region amino acid sequences and segments of sequences in antibodies of different specificities. Relative contributions of VH and VL genes, minigenes, and complementarity-determining regions to binding of antibody-combining sites. J Immunol. 1991;147:1709–19.
- Vandyk L, Meek K. Assembly of IgH CDR3: mechanism, regulation, and influence on antibody diversity. Int Rev Immunol. 1992;8:123–33.
- Jespers L, Schon O, Famm K, Winter G. Aggregation-resistant domain antibodies selected on phage by heat denaturation. Nat Biotechnol. 2004;22:1161–5.
- Pechmann S, Levy ED, Tartaglia GG, Vendruscolo M. Physicochemical principles that regulate the competition between functional and dysfunctional association of proteins. Proc Natl Acad Sci USA. 2009;106:10159–64.
- Saphire EO, Parren PWHI, Pantophlet R, Zwick MB, Morris GM, Rudd PM *et al.* Crystal structure of a neutralizing human IgG against HIV-1: a template for vaccine design. Science. 2001;293:1155–9.
- Trovato A, Chiti F, Maritan A, Seno F. Insight into the structure of amyloid fibrils from the analysis of globular proteins. PLoS Comput Biol. 2006;2:e170.
- Tian J, Wu N, Guo J, Fan Y. Prediction of amyloid fibril-forming segments based on a support vector machine. BMC Bioinformatics. 2009;10:S45.
- Caflisch A. Computational models for the prediction of polypeptide aggregation propensity. Curr Opin Chem Biol. 2006;10:437–44.
- Cerdà-Costa N, Esteras-Chopo A, Avilés FX, Serrano L, Villegas V. Early kinetics of amyloid fibril formation reveals conformational reorganisation of initial aggregates. J Mol Biol. 2007;366:1351–63.
- Chen Y, Dokholyan NV. Natural selection against protein aggregation on self-interacting and essential proteins in yeast, fly, and worm. Mol Biol Evol. 2008;25:1530–3.
- 64. Linding R, Schymkowitz J, Rousseau F, Diella F, Serrano L. A comparative study of the relationship between protein structure

and  $\beta$ -aggregation in globular and intrinsically disordered proteins. J Mol Biol. 2004;342:345–53.

- Rousseau F, Serrano L, Schymkowitz JWH. How evolutionary pressure against protein aggregation shaped chaperone specificity. J Mol Biol. 2006;355:1037–47.
- Reumers J, Maurer-Stroh S, Schymkowitz J, Rousseau F. Protein sequences encode safeguards against aggregation. Hum Mutat. 2009;30:431–7.
- Cecchini M, Rao F, Seeber M, Caflisch A. Replica exchange molecular dynamics simulations of amyloid peptide aggregation. J Chem Phys. 2004;121:10748–56.
- Cellmer T, Bratko D, Prausnitz JM, Blanch HW. Protein aggregation *in silico*. Trends Biotechnol. 2007;25:254– 61.
- Nguyen HD, Hall CK. Spontaneous fibril formation by polyalanines: discontinuous molecular dynamics simulations. J Am Chem Soc. 2006;128:1890–901.
- Tuite MF. Yeast prions and their prion-forming domain. Cell. 2000;100:289–92.
- Chen SM, Berthelier V, Hamilton JB, O'Nuallain B, Wetzel R. Amyloid-like features of polyglutamine aggregates and their assembly kinetics. Biochemistry. 2002;41:7391–9.

- 72. Pieri L, Bucciantini M, Nosi D, Formigli L, Savistchenko J, Melki R *et al.* The yeast prion Ure2p native-like assemblies are toxic to mammalian cells regardless of their aggregation state. J Biol Chem. 2006;281:15337–44.
- Hamada D, Tanaka T, Tartaglia GG, Pawar A, Vendruscolo M, Kawamura M. Competition between folding, native-state dimerisation and amyloid aggregation in β-lactoglobulin. J Mol Biol. 2009;386:878–90.
- Lehrman SR, Tuls JL, Havel HA, Haskell RJ, Putnam SD, Tomich CSC. Site-directed mutagenesis to probe protein folding: evidence that the formation and aggregation of a bovine growth hormone folding intermediate are dissociable processes. Biochemistry. 1991;30:5777–84.
- Baden EM, Randles EG, Aboagye AK, Thompson JR, Ramirez-Alvarado M. Structural insights into the role of mutations in amyloidogenesis. J Biol Chem. 2008;283:30950–6.
- Azriel R, Gazit E. Analysis of the minimal amyloid-forming fragment of the islet amyloid polypeptide. An experimental support for the key role of the phenylalanine residue in amyloid Formation. J Biol Chem. 2001;276:34156–61.
- Koide A, Gilbreth RN, Esaki K, Tereshko V, Koide S. Highaffinity single-domain binding proteins with a binary-code interface. Proc Natl Acad Sci USA. 2007;104:6632–7.