#### PROTEINS: Structure, Function, and Genetics 21:1-10 (1995)

## **PREDICTION REPORT**

# Predicted Secondary and Supersecondary Structure for the Serine–Threonine-Specific Protein Phosphatase Family

Thomas F. Jenny, Dietlind L. Gerloff, Mark A. Cohen, and Steven A. Benner Department of Chemistry, Swiss Federal Institute of Technology (E.T.H.), CH-8092 Zurich, Switzerland

ABSTRACT A bona fide consensus prediction for the secondary and supersecondary structure of the serine-threonine specific protein phosphatases is presented. The prediction includes assignments of active site segments, an internal helix, and a region of possible  $3_{10}$  helical structure. An experimental structure for a member of this family of proteins should appear shortly, allowing this prediction to be evaluated. © 1995 Wiley-Liss, Inc.

Key words: protein structure prediction, protein phosphatase, evolution

#### INTRODUCTION

Four developments have characterized recent work to develop useful approaches for converting sequence data into models of the conformation of a polypeptide chain. First, an explosive growth of sequence databases has permitted conformational models to be built from alignments of homologous sequences rather than from single sequences.<sup>1-4</sup> Second, the explosive growth of sequence databases has enabled the development of new, often highly detailed, models of how protein sequences diverge under functional constraints.<sup>5-8</sup> These have allowed the development of prediction tools that transcend simple averaging of classical predictions over a set of aligned homologous protein sequences, in particular, those that extract conformational information from patterns of conservation and variation within a multiple alignment.9-11 The combination of abundant sequence data and detailed evolutionary models has underlain a third development, where tertiary structural information is explicitly introduced early in a structure prediction, providing a way free of the vicious cycle that arises because secondary structure is examined before tertiary structure is predicted, but tertiary interactions are the strongest determinants of secondary structure. Fourth, bona fide predictions, those that are announced before an experimental structure is available, have again had their value recognized as a key part of an effort to

develop useful prediction tools, especially as a way to identify and avoid biases that inevitably compromise methods tested using a database known at the time the tools were developed.<sup>12</sup>

The primary disadvantage of bona fide predictions as research tools is that such predictions are made one at a time, and must await confirmation by subsequently determined experimental structures. Thus, it is difficult to generate rapidly a large number of predictions that can permit a statistically significant test of a prediction method. This implies that efforts to use bona fide predictions as a research tool must be sustained over a substantial period of time to allow accumulation of many predictions. Especially important in these efforts are venues (such as this) where such predictions can be published, as these permit a record to be established of the scope and success of prediction methods, creating an accountability that itself serves the research effort.

To date, over a dozen bona fide predictions have been made in four laboratories using tools that include an analysis of conservation and variation in homologous protein sequences. With the recent prediction of the structure of protein tyrosine phosphatase<sup>13</sup> by Livingston and Barton,<sup>14</sup> eight of these can now be evaluated using one or more subsequently determined experimental structures.15 While no finite number of successful predictions can "prove" a method, these predictions have shown that these methods are transferrable from laboratory to laboratory, can provide some remarkably accurate secondary structure predictions, and can be the first step in the generation of tertiary structural models from secondary structural models. Particularly noteworthy is the first prediction contest that placed these tools in competition with the most advanced neural networks and classical methods averaged

Received July 12, 1994; revision accepted September 19, 1994.

Address reprint requests to Dr. Steven A. Benner, Department of Chemistry, E.T.H., CH-8092 Zurich, Switzerland.

over a set of aligned protein sequences.<sup>16</sup> In this contest, predictions based on an explicit analysis of patterns of conservation and variation assigned more residues correctly than either the neural networks or classical tools. More important, the number of misassignments that confused  $\alpha$ -helices and  $\beta$ -strands in the prediction was vanishingly small, suggesting that the new tools yield secondary structural models especially suited as the starting point for tertiary structural modeling.<sup>17</sup> Again, however, this was a single case, and many more prediction contests must be arranged before the relative merits of different methods can be evaluated in any general way.

Recently, prediction efforts have focused on proteins and domains involved in signal transduction, including the Src homology 1 (SH1) domain, a protein kinase,<sup>18</sup> the Src homology 2 (SH2) domain,<sup>19</sup> a protein unit that binds peptides containing phosphotyrosine, the Src homology 3 (SH3) domain,<sup>20,21</sup> a protein unit that may bind proline-rich peptide sequences, the protein tyrosine phosphatases,<sup>14</sup> and the pleckstrin homology (PH) domain,<sup>22,23</sup> a domain with unknown function.<sup>24,25</sup> We turn now to the protein serine-threonine phosphatases,<sup>26</sup> a family where no experimental structure is known, but where an experimental structure may be imminent.<sup>27</sup> We report here a predicted consensus secondary structure for this protein family.

#### METHODS

Sequences of protein serine phosphatases were extracted from entries in SwissProt 27 and a multiple alignment built by DARWIN.<sup>28</sup> The multiple alignment was then adjusted by hand (Fig. 1). In particular, the nonaligning segments at the beginning and end of the multiple alignment, presumably representing noncore regions, were removed, and the insertion in human calcium-dependent calcineurin A (accession number P16298; sequence p, between alignment positions 073 and 074), presumably representing an intron-derived sequence, was removed. Surface and interior residues were then assigned by an automated computer program, following a design described in detail elsewhere.8 The multiple alignment was then parsed (separated into segments that form independent secondary structure) by analysis of gaps in the multiple alignment, identifying conserved Pro and Gly residues adjacent to surface using procedures, and identifying parsing strings, using heuristics described elsewhere.<sup>8,9,16,18</sup> Positions that have conserved functionality in the context of strings of consecutive or nearly consecutive active site assignments were assigned to the active site, as discussed elsewhere.9,16,18 Many of the automated computer programs used in this work are available on a server via electronic mail at the address cbrg@inf.ethz.ch (send a one word message "help" for instructions).

Secondary structures were then assigned to individual parsed segments of the multiple alignment from patterns of conservation and variation in the various assignments using an automated computer program that detects periodicity in these patterns. A 3.6 residue periodicity in interior and surface assignments was designated a surface helix, while alternating periodicity in these assignments, was designated a surface beta strand, as discussed elsewhere.<sup>9,16,18</sup> Short ( $3 \le 6$  residues) segments of interior assignments were designated as β-strands, while heuristics designed to distinguish between internal helices and internal strands (see below) were applied to longer segments. Segments of consecutive or nearly consecutive residues assigned near the active site assignments were assigned as "active site regions." In these regions, patterns of amino acid substitution are dominated by constraints relating to catalytic function, which often obscure patterns that reveal secondary structure.

The secondary structures predicted here are presented in a way that facilitates their use as the starting point for tertiary structure modeling. Important in this presentation is recent work that defines the limits to which a consensus secondary structural model built from a family of proteins can represent the conformation of any individual member of the family.<sup>12,29,30</sup> In building tertiary structural models from predicted secondary structure, the  $\phi$  and  $\psi$  angles in segments assigned as helices or strands are held rigidly, while those in the coil segments are permitted to vary. Therefore, the most useful secondary structure model is presumably the one that assigns helical or strand conformations to the cores of these secondary structural units, those likely to be found in all members of a protein family, and leave flexible those regions that are not likely to be found in all family members.

Supersecondary structure, defined here to describe the relative orientation of consecutive secondary structure units, was predicted based on orientation of predicted secondary structure units with respect to active site assignments, and searching for compensatory covariation, as described elsewhere.<sup>18</sup>

#### **RESULTS AND DISCUSSION**

The secondary structure prediction is summarized in Table I, based on the multiple alignment, surface, interior, parsing and active site assignments recorded in Figure 1. Most of the predicted secondary structural elements were obtained by automated procedures that implement the prediction heuristics outlined in 1989 and 1991.<sup>9,18</sup> Certain of the secondary structure segments are canonical,<sup>18</sup> and should be highly reliable. In particular, helices (003–019), (119–130), and (236–243) are reliably assigned (Fig. 2), as are strands (059–062), (090–093), (135–139), (143–146), (175–180), and (228–232). A helix can also be assigned to segment (101–116), provided that a weak surface assignment at position 107 is accepted.

The remaining assignments of secondary structure were not routine. For example, the interior and surface assignments made for segment (154–166) display a perfect 3.0 residue periodicity. This is expected either for an  $\alpha$ -helix whose contacts with the remainder of the protein precess around the helix axis, or for a 3<sub>10</sub> helix. 3<sub>10</sub> helices are rarer than  $\alpha$ -helices, but are not unknown in protein structures.<sup>31,32</sup> This is, however, the first time such patterns have been observed in the context of a bona fide prediction. We have listed the helix as being of the 3<sub>10</sub> type to encourage more effort to investigate patterns of divergent evolution in such "nonstandard" secondary structures.

Internal helices are also difficult to find using methods that assign secondary structure based on periodicity in interior and surface assignments. An internal helix was missed in the protein kinase prediction,<sup>18</sup> while another was correctly identified in the hemorrhagic metalloprotease family.16 In the protein phosphatases, an internal helix is assigned to positions 075-086. This assignment is based on two criteria: (1) the length of consecutive interior assignments compared with the overall size of the protein, and (2) the pairing of the helix with a  $\beta$ -strand following it (090-093), where the short parsing element (087-089) and the active site assignments within positions 063-074 and 094-098 require the two structural units to be antiparallel and roughly equivalent in length. This is noteworthy as an example where the assignment of secondary structure strongly relies on a tertiary structural hypothesis, and it will be interesting to learn whether it is correct.

The most problematic assignments for any method for predicting secondary structure are those made for segments near the active site. Parameters derived from a statistical treatment of protein sequences as a whole are the least representative in these regions. Further, near the active site, patterns of variation and conservation are dominated by the demands of catalytic function, which often obscure patterns that might indicate secondary structure type. Previous prediction efforts have shown how problematic these regions are in other protein families.<sup>29</sup> and similar regions are encountered in the protein phosphatase family. For example, the segment 029-050 includes a β-strand assigned to segment (030-034) and a short  $\alpha$ -helix assigned to segment (043-049) separated by an active site region (035-042) with poorly defined secondary structure. The  $\alpha$ -helix might be extended into the active site region. However, the conservation evidently arising from catalytic function precludes reliable tertiary structural assignments in this region that might secure this extension.

Further, the helix assigned to segment (246-262)

contains a conserved tripeptide RxH that is plausibly (but not definitely, see below) placed at the active site. The segment displays convincing 3.6 residue periodicity (Fig. 2) if residue 254 is assigned to the surface. To observe this periodicity requires, however, assignment of a conserved R (251) to the surface and a conserved H (253) and a conserved G (259) to the inside. Further, the DG element at positions (258-259) is a weak parsing element. Thus, a second, weaker, assignment separates this segment into two shorter elements separated by an active site coil. In tertiary structure modeling, this alternative assignment must be considered.

Finally, the final segment (274-299) contains a conserved SAPNY that is also plausibly placed near the active site. The segment (274-278) might canonically be assigned as a  $\beta$ -strand, as is segment 290–295. This leaves, however, segment 279–289 problematic. Dipeptide parses at positions (281–282), (284–285), (288–289), and (291–292) suggest that this segment is a coil. However, the first is involved in a putative active site assignment, making the assignment lower in reliability.

While active site assignments are sources of ambiguity when assigning secondary structure, they make key contributions when building a tertiary structural model (Fig. 3). Several long, conserved, functionalized peptide segments are convincingly assigned to the active site. First, the segments (035-042), (063-074), (094-098), and (174-152) are clearly present in the active site. The isolated conserved functionalized residues at positions 83, 111, 177, 182, and 226 are not. The segment (279-283) is also assigned to the active site, even though it contains an unusual pattern of functionalization that suggests a noncatalytic role. Finally, the segment 251-253 is also assigned at the active site, primarily based on the conservation of His-253.

Based on these active site assignments and patterns of compensatory covariation, the protein kinase structure can be divided into several supersecondary structural elements (Fig. 3). For example, both strand (030-034) and strand (059-062) end at the active site. They are joined by helix (043-049), with the loop connecting helix (043-049) and the following strand not lying at the active site. This forces these three secondary structural elements into a  $\beta - \alpha - \beta$  segment where the two  $\beta$ -strands are part of a parallel  $\beta$ -sheet. With only a few exceptions, this supersecondary structural element is right handed. Further, refinement of the secondary structural model in light of this tertiary structural modeling suggests that the helix (043-049) probably includes additional amino acids at the end, most probably residues (040-042) and (050). Together, these considerations support a good model of the conformation of these 35 amino acids.

Similarly, internal helix (075-086) starts near the active site, while the following strand (090-093)

alignment indicate sequences that are fragments. Secondary structure: A, α-helix; B, β-strand; \$, ac-tive site segment; 3, 3<sub>10</sub> helix. a–(P16299) protein phosphatase 2B (calmodulin-dependent calcineurin A2), *Homo sapiens*. b–(P29349) protein-prosina-phosphatase (EC 3.1.3.48), *Drosophila melanogaster*. c–(P23635) protein phosphatase PP2A-1, *Schi*-P—(P16298) protein phosphatase 2B (calmodulin-dependent calcineurin A1), *Horno sapiens*. q—(P20652; P12816) protein phosphatase 2B (calmodulin-dependent calcineurin A), rat and romyces cerevisiae. —(P05323; P13197) protein phosphatase PP2A-α, *Homo sapiens*, rabbit, pig, and boj--(P13353) rat protein phosphatase PP2A-α, rat k--(P11082) protein phosphatase PP2A-β, *Homo* Positions in the alignment are numbered consecu-tively. Target position numbers refer to the human calcineurin A2 sequence. Conserved residues bearing functional groups are given in upper case letters; others in lower case letters. Numbers in surm—(P11611) protein phosphatase PP2A-8, rabbit, n—(P23778) protein phosphatase PP2A (Irag ment), rape. o-(P23696) protein phosphatase PP2A, Dro chosen to be Giy or Pro adjacent to Giy, Pro, Ser, Asp, or Asn. Tripeptide and higher parsing strings are consecutive (Giy Pro Ser Asp Asn) residues. Lower case letters in the sequences indicate residues not aligned by DARWIN. Lower case x in the zosaccharomyces pombe. (---(P23595) protein phosphatase PP2A-2, Sach---(P32838) protein phosphatase PP2A. Saccha-I—(P11493) protein phosphatase PP2A-β, (Irag Fig. 1. Consensus secondary structure predic-tion for the protein serine-threonine phosphatases. face and interior assignments increase qualitatively with increasing strength of the assignment. Index numbers in the multiple alignment refer to se-quences from SwissProt 27 as listed below. Parsing strings are underlined, with dipeptide parses e--(P23636) protein phosphatase PP2A-2, Schi g--(P32345) yeast protein phosphatase, Sacchacharomyces cerevisiae. charomyces cerevisiae. sophila melanogaster. romyces cerevisiae. sapiens, ral. ment), pig. mouse. -----~~~ ARG COCO ļ 15 88 555 255 0000 0000 GCCCC NNNN SEE KKKK 1 ]]] 13 1 H N N AMONA VIV HHH EEED Raao 301. 6550 PPPP 5555 **JAR** 0000 1/111 500 LLLL. TITT 0000 11HI 00TR NCK N 2000 SEEE ş 2 722 RERRERERE EVERTEREE E HEREFORMENT HERFORMENT HERFORMENT HERFORMENT HERFORMENT HERFORMENT HERFORMENT HERFORFORMENT HERFORFORFORMENT HERFORFORFORFORFO 000000000000 TREFFERENCE T 5 505050505050 1 333333333533 actanactana **JECEEGEGECE** MANANANANANAN AAAAAAAAAA WHEN HEN IIIIIAN PPPPPPPPPP 11111111111 11111111111 DENAMAAATI TELLELELLE RARRAR **JAJAJAJA** IVLLELELIN ചദാ ¢ 194 
 0
 CONTRACTOR

 1
 CONTRACTOR
 </tr HEDEEEELMDEEDE V \$\$ DDDDDDDDDDDDDD H HR ETTVC00015555F5 D G0,G0000005555F5 D α NNKNSNNANNNNANN NNKNSNNANNNNANN REARCHARKERERE GUIGGGGGGGGGGGGG MED OGGS SOCGAGAS SES TUTHOUT SEENT **PAME** 0000000000022 AVAALI ITT. \*\*\*\*\*\*\*\*\*\*\*\*\* H Ľ 2 8 SSN SSN 28 읪 2 ÷ 3 8 85 0.23 0.19 0.23 0,09 0.09 0.25 0.39 0.25 0.30 0.19 1.35 0.29 1.37 1.35 9.46 0.09 0.09 0.70 0.09 09.3 2.60 9.94 0.10 2.60 0.10 2.55 2.60 0.07 2.69 

 0.11
 1.11

 1.13
 1.11

 1.13
 1.11

 1.13
 1.11

 1.13
 1.11

 1.13
 1.11

 1.13
 1.11

 1.13
 1.11

 1.13
 1.11

 1.14
 1.11

 1.15
 1.11

 1.11
 1.11

 1.11
 1.11

 1.11
 1.11

 1.11
 1.11

 1.11
 1.11

 1.11
 1.11

 1.11
 1.11

 1.11
 1.11

 1.11
 1.11

 1.11
 1.11

 1.11
 1.11

 1.11
 1.11

 1.11
 1.11

 1.11
 1.11

 1.11
 1.11

 1.11
 1.11

 1.11
 1.11

 1.11
 1.11

 1.11
 1.11

 1.11
 1.11

 1.11
 1.11

 1.11
 1.11

 1.11
 1.11

 1.11 Kultiple Alignment Sec JXuirBCAGDEyvkLhidnolmkjiec IHgweispas Struc VVVF ERED RAM 5555 FFEF FFEF lvvv eddd eeee EEEE VVVV GOOG age, ordorargoorangiae, a lastaranargaa, aana, araa YYYY VVVV DDDD DDDD RBRR RBRR LLLL LLLL 1111 ARRI 222 2 200 DDGG MARCH IN THE PARTY OF THE PARTY 200 KKFk FFFFFFFF F Kkrrrrrrkn r GGGGGGGGGGGGG A SV222222222 111111111111 14111111111 adoooooda adoooooda COCCCCCCCCCCC ( 3333333333333 าหารกรรรรษ HANDARD DANAH SSSTNWNSSSS EEEEEEEEE RATTTSSMO **THKREKKKKKL MANANANA** TYYYYYYYYY LLLLLLLLLL FFFFFFFFFFFFFFFFF 000000000 WWWWWWW EEEEEEEE AAARFURAKKEI LLANDSSHUND \*\*\*\*\* REALEDED 11 liliiiilillillill [ ] se ksselsetsetsen 5 25 ak edterforsten 5 25 ak edhospyrstaann v D2 ak edhospyrstaann v V2 ak ed LICHTOLLELELLIN H KEEEEEEEEEDEEDDTK O eftererererer t a a ETTITITITY VITY V ETTYYGGGOORHMOL A ИСССССССССАЧИСС С 47ИSTLLLAAMMIY F 97ИККИКАРЖЕЕКИ К 3 AGAGAGAGAGAGAGA ه, REFERENCESSER SAAA555A5x55AA ryffefferryfref LULLULLULLULL \*\*\*\*\* LLLLLLLLLLLLLL XXXXXXXXXXXXXXXX RRAURBRARKARA 

2 61 0 91 0.14 1.31 21.35 0.46 Sur-face 2 8 8 7 7 8 8 7 7 8 8 7 5.25 0.13 0.23 0.96 0.09 Inteş;; 0.30 2.60 2.60 2.60 0.42 0.10 2.60 2.2 2.60 0..30 0.17 3 Conser ved ėο = 00 ۵ Position Align Targ 1945 

118822833

žJŁ

ដន់មិនិង

12565853F3

23828

1919

5

759262

닅븝

		`	b f (r < 3287) protein phosphatase 2B (calcineurin)		5	dependent calcineurin A) rat	I(P14747) protein phosehotana an	A2) Sarrhamming - rushing as (Calcineurin	II ID12641	3	Saccharomyces pombe	ł			 >		romyces cerevisiae	X(P23090)		Saccharomyres nomber 1 Juni20-	in (Dooto a) and the pulling	y-(rzd/34) protein phosphalase Toman		-	Z(P32598) molein	,		•	A-(PU8129: P22803: Dancrai		priatase PP1-0. Homo caniano	ď	sphatase	Sonhila melanomica	Coprime meranogaster.	C(P08128) protoio atomic -	Construction protein prosphatase PP1-8, rabbit	T	The second	Sis Inaliana		(r <3/ //) prolein phosphatace pp1 (		idhe.	E/P306E41	Time(r 20034) Diotein nhosobalasa nha a sa	him mid the production of the ASDEROIL		(i)		H-(P33179) protein abarret	A Providing Providing PP-X Home	saplens		The root protein phosphatase pp_Y methic	J—(P11612) protain absorb 11 18001.	V Proventi prosphatase PP-Y Drosonh.	lia melanogaster	K(P26570: 000070)	Contraction Goustal protein phosphatace pp.	21. Saccharomyrae coronician	Provide a cardination of the car	L-1/132945) prolein phosphatace CAL C Contra	Inmures certaining	and the second s										-	_								_		-		-		
		A MANANANAA AAAA CCC	FYFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF	FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF	DGGGGGGGGGSSNG C COCCOLLER FFFF	YF KAPPRAARAAREKK R ONONONON COUL	AK VDDDDEEEDDDSSVR S DDDDDDDDDD	AA IVVVVVKKKKVVAN V IVIIIIIII	VAVA VAVAVAVAVA V 25555555 V VIVVVVVVVVV CV	HX DDDDDVVXTDDDDVXXKNLE EENEEEEEE AADR CCCC	CODE DEEC	· · · · · · · · · · · · · · · · · · ·	T HUNNHANNAN MALTA LILL	AAEH 0000	H TTSTANAAAAA AAKV KNNN	ALCON NAWN NAWARAANAANAANA N TTTTTTTTTTTT	D DDSGGGGGGGGGGGGG	H LITTLE A		a south TTTTSD	×.	IIIVVVVVII I	Y MSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS	a boobnorsee class Lill	A THE REAL PARK	ANN MAN MANANAN ANN	N RHRAHMARAMA KANA HUSA	0 0000000000000000000000000000000000000			,	CODECCERTEDDEE N HERERARMANAM YOUN DOOD	WHUNDODDAADD & BEDEKEEEEE EFEF		•	Y YIYYYYYFY YVYF	PERFERENCES O SANNANNANN KANK			WE FREEDERSON TO SSACCCCCAT HRIH TYYY		MGSSSAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA			MUGHURRARY C OOARABAGH TTO									Line and the second second	ATAA MAMAAAAA A ATTITITITITITI	AVVVVVVVVVVV T VVCVVVVVVVVVVVVVVVVVVVVVV	TITITITITITI T TTTTTTTTTTTTTTTTTTTTTTT	V 11111111111	FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF	SESSESSESSESSES & CONTRACTOR NAME FFF	3555 S555	DEPENDENCE A AUMANAAA AAAA AAAA	MODACHARTER P. PPEPEPEPEPE PPP	MARTING AND AND A STANDARD AND AND AND AND AND AND AND AND AND AN					THE PROPERTY AND	C CCCCCCCCCCCC CCCC	NOUDDDDDDDDDDD R KOODDDDDDDDD R KOODDDDDDDDDD	V MORPHERE CONTRACTOR	3	K 0000000000	A ANAAAAAAA AAAA		,	I IIILIIIIII IIW WWW	L PHLIPSPORTH LLIN TITLE	Z EEEEEEEEE	WWWWWWWWW & WARFERS	A DOGROSSIC	•	AN REALFACTERIZET & REODDODDD EEDE NNNN						
		-		2.60	0.19	0.45	1.20	0.47	1.26	0.96	2.60	0 16	3.46	0.90	91.1	-			1.68	9.20	1.79	0.40			0,45		0.11	, co			0.03	0.44			2.60		0.19		0.18	0.23	CL 0		1.39	0.70		0.15	0.15	•.		0.09			1.26	0.42	0.25					;	7.60			2.60		2 :	9.17	0.16	2.37	2			.06	0.45		16.4	2.60	60				0.90	2.70	1.70					•	
																							æ	4		<b>T</b>							0																	-				•	Ŭ				" 			م	z		,		Ð		2		*		•	-			~	~		•	5									•
																																																				~			216	3	-	15	16	13		<b>D</b> 1	616	120	321	5		3	124		326			59	329		2	-	32	66			, n	9						
	265	266 267	268	269	271	272	672	274	275	9/7		0.7		0.07	187	282	283	284	285		987	C87	288	289	195		167	292	191			295	296	247		. B6Z	299	000	191		302	IDE	101	• • •	50	No.		100		808	ŝ	=			•	~						•						÷ 4		Ň					в			2	Ξ.	2	- 2			m -						
	22B 265				234 271		ETS 3ES	237 274	272 822	9/7 603		747 278			187 667		246 283	247 284	202 812					252 289			167 677	255 292	256 293	257 24		567 867	259 296	260 247		862 142	262 299	263 300	264 101		205 202	266 303	267 304		268 305	269 200		TAC 015		R05 1.3	m	273 110	27.4	i.	3	216	5	5.15	279 3	280 3	281				284 31	285	386			288 325	289 33	290	• •	67	292						100 162									
	<	A. 229	107 V	217 V	× 234	A 235	9CZ Y	- C		53		2		22		2 2	546	247	2 T	5 249	22		, 55I	222	253	A 10		52	A 256	A 257			523	2 260				52	A 264			566		-	, 26B	592		FOC 012		-	E 217 Y	× 273	274	1.1	<	*	1 A 277		1 A 279 3	4	283		7 .	ē	284					88 788		a 240			262	1.6Z		51									<b>)</b> U		•	
	<	LLUT LLLL A' 230	6 TTGGLGLGLGL LLLC YYYY A 231 b bbubchtatt, rot	LILM MWWW A 232	AACC AVVV A 234	A 235	XXXX XXXX X X X X X X X X X X X X X X			LEEP PEPP	HUGEEEEENO DDDA KSSS 741	RRRK TTTT 242	IIIIIIIIII IIII TTTT P	TTT FFEF 6 211		TTTV 1411 P 245	RRRRRRR PDF PPF 246	247	5 T	5 249	ИННИЧИНИИ ИНИИ В 250	EEEEEEEE EEEE EEFE 6 20	55555555555 cere root 251		COLORANCE REAL RAPE 251	1000000000 0000 HHHH 0000 000000000	IIII LILL I	522	22 V 111 111	UCING EEEE A 257		YYY FFFF A 29 4	259	2002 200 2		YYY KKKK		Lete reer 263	LUCTORDOOD TO BEER A 264	ALL A COLC CONC CCCC Y CALL	LLVL XKKK A 253 3	DPDN 111 266		XKKK KKKK A	, 26B	269		270	SEMNNINNSN SSNT FOOD		TAT TELEVEL AND TATE TELEVEL Y TIS 3	NEWNWWWWW THET REAL A 273		2002	V III muu	KARK DEEE A		TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT		EEEO DEEE A	LIVY AAAA 201		Date of the A total and a set		TTTP CSSS 284	FPLLLLLLLFL IV 285 1		ddd Leve		SSGA AAAA B 28B	289 AMA AMA 8 289	TILL LLT. a 240			HANNA DIGG	CGNG 0000 241	XEST DOOD		567 B 3342 1111	FFFL LLLL B 296	CCCC CCCC B. 297	WWW WWW B 298	HNNH HSHM C 200			Lukkhahata I.I.I. I.I.I.	\$555	EP.	MSSSSSSSS SSDE EEEE	
	T TTTTTTTTT TTST CCCC A	M SSTTTTTTSS LLET LLLL A 229	G G PARTAGARATANA CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONT	I VVVVVVVVI LILLM MWWW A 232	V ANGANANANA AACC AVVV A 234	W MAULULITHE LLYL LULL A 235	L WWWWWARKA KKKK A 236	R REREVERENCE REPORTED 1 237	BEZ AAAA AAAA AAAAAAAAAAAAAAAAAAAAAAAAAA	2 PPPRPPRARP REP PPPP	5 NHOEEEEENO DDDA KSS5 741	R REREARINGER RARK TITT 742	I IIIIIIIII IIII IIII P	H TTTTTTTTTTTTTFFFF	L HIHITITI LICE THE COMPANY	L ERLEILLEL ITTV LILL B 245	R REREARDS SOON NOT 246	247	5 T	5 249	И ИННИЧЕНИИ В ИНИИ НЕИК 2 250	E EEEEEEEE EEEE EEEE	5 55555555555 FOT TOOL 251		CONTRACTOR RANGE RANGE 251	A 000000000 NHHH 0000 00000000 N		7 TTTTTTTTT ++++ ++++ 255	0 000000000 111 1 1 256	C MARCANANAN UQIKO EEEE A 257	AT A YAYA WWW WWWWWWWW C	I TYTYTYYYYY YYYY FFFF A	G GGGGGGGG GDGC TTTT 259	FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF		Y YYYYYYYYYY YYYY KKKK		E EELERERER FFFF FFFF FFFF 263	CONTRACTOR DECE A 264	A THE A COLO CONC CCCC Y	L LULLULLULU LUV XKKK X 103 3	N RRRRRRRRR DODN 111 266			, T YTYYYYYYY YYYY YYYY . 268 J	269		A 14401566666 6666	A SANNANANSA SSNT TCCC	S AAAAAAAAA mirita waxaa ahaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa	T T T T T T T T T T T T T T T T T T T	A NHANNANNAN TTET RARH A 273	WWW WWW A	N KIMMINIMMINI VILLA	A TILI THAN ANYWARAN O	A MAYAYAWAA RARK DEEE A	YYYY THYYYYY	LITTLEFFFF COCC	JOTT TTTTTTTTT	D UDUDDDDND EEEO DEEE A	I LUCULULUL IIVY AAAA 201					L FFLLLLLFF, I	V PPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP	T UTT TT T		288 AAAA B 288	CANNANNA MANA AMAA B 289	I LULULULUL IIII LULU a 240	I WWWWWII III III III	D DDDDDDDD	WNWN DIGO 337000000000	D MNSCGGGGDD GGNG 0000 211	E KKEODODAR KKEN MANA		S67 6 JAA TIT IIII ALAS J	C CONCESSION OF A LEFT LLLL B 296	C LUCCOCCCC COCC CCCC B 297	V LIGHTLELLEL WWW WWW R 298	A MARMARMANAN HANA HANA HANA C	6 6666666757 Cree cree	<u>6.0000000000 0000</u>		\$555	EP.	ğ	
	TTST CCCC A	L M SSTTTTTTSS LLET LELE A 229	G G PARTAGARATANA CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONT	I VVVVVVVVI LILLM MWWW A 232	YYYYYYYYYYYY AAGAAAAAAA AACC AVVV A 234	KKKKKKKKKKKK F FERELELELENF LEVE LELE A 235	L WWWWWARKA KKKK A 236	KKKKKKKKKSSKK R REGURREGER BERK 111 A 237	BEZ AAAA AAA AAAAAAAAAAAAAAAAAAAAAAAAAAA	TETERSPERSED POPERPERSED PEPP PEPP	PERCEPTER SEEED 5 NHOREELEENO DDDA X555 741	FEETENWANKTING REREARDER REAK TITT 242	I IIIIIIIII IIII IIII PUI		PULIVOLDELEDIN L ILTITITITI LELL THE CONTRACT	UPDEPERTURE LEALENEEDEL ITTV ATT 245	WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW	247	5 T	5 ANNUN MANN MANN SAMANA SAMANA	EFFERERERERERERERERERERERERERERERERERER	E EEEEEEEE EEEE EEEE	Crucccccsbcccs 5 55555555555555 Ford 251	AMAAMAAASSAA P PPDPPPPPPPP 3373 CCCC 252	SSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS	V HHHH 0000 00000000 N HHHH V		255 222 222 222 222 222 222 222 222 222	KRRRRRRRRRRR A CONCOMMENT A 256	C MARCANANAN UQIKO EEEE A 257	AVYVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVV	THE TRANSPORT TYPEST AND THE TYPE TO THE TYPE TH	HOUNDOGGGGGGG G GGGGGGGGGG GGGG GGGG TTTT 259	EFFEFEFEFEFEFEFEFEFEFEFEFEFEFEFEFEFEFE		THE PROPERTY A TYPYTYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY		ANDERDERERE RECERCERERE FOR FOR FOR 263	LUCTORDOOD TO BEER A 264	NERREARE COLOCIC CONC CCCC A COLOCICIÓN ANALANA	DESCRIPTION OF A CONTRACT OF A	PURKHURRARRARRA N RERARRARRAR DEEN III 266		T TATARAN KAKA KAKA KKA T	, T YTYYYYYYY YYYY YYYY . 268 J	269		270	The same and a substantian source and a substantian and a substant	VIIIIIIIVUVVIS SARAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA		273 A PROVININAN TTET REAL A 273	WWWWWWW V WWWWWWW	NONTRACTICAL A NUMBER OF A DAMAGED AND A	HKKKKKKKKKKKK O KKKKKKKKKE SOUL	A MAYAYAWAA RARK DEEE A	FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF	COCC - COCC	DDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD	CONTRACTOR DECE A	VULLEVELTTY & ELECTERELE I IVY AAAA 201			MOCCCCCCCCTT Y YYYYYYYYY HUDD 10000 281		PARTELLINGLE L FFLETLING		Addd Leer Thirty I TAWMAIIAIIMA		ANAMAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	LUTITITITIE AND AND AND AND ANA AAAA B 289	VIIIIIIII I LULIULULULULULULULULULUU	I WWWWWII III III III	CODDDDDDDDEEAO D NDODDDDDDD 2011 NLUL B	WNWN DIGO 337000000000	PERMANANA D' MNSCRCCCCDD CCMC 0000 241	TANANARARARAR E KACOOOODAR XEER DOOL		FEFFERLLLEF & PEPPERSTER ALL FYFE 9 295	C CONCESSION OF A LEFT LLLL B 296	CHEMICATION C LUCCOCCCC CCCC B 297	WINGSTRUCTURE OF A LIGHTLEFELL WAVY VAVV R 298	ОСССООЛОНИИНИИНИИ И ИМИНИИИ И ИМИНИИ И И И И И И	6 6666666757 Cree cree	G.GGGGGGGGGGG COCC	1111111111111111	\$555	PEPEPEPEPEPEPEPEPEPEPEPEPEPEPEP	ğ	
	T TTTTTTTTT TTST CCCC A	M SSTTTTTTSS LLET LLLL A 229	LEGENERALELELELE VYLERELELELE LELE YYYY A 231	FFLGLLGLELLFL I VVVVVVVVI LILM MMMM & 232	YYYYYYYYYYYY AAGAAAAAAA AACC AVVV A 234	KKKKKKKKKKKK F FERELELELENF LEVE LELE A 235	The VIIIIIIVIILLUI L'UNUNUNUTI IIIVII	KKKKKKKKKSSKK R REGURREGER BERK 111 A 237	BEZ AAAA AAA AAAAAAAAAAAAAAAAAAAAAAAAAAA	DE LETTERATUREPERK 2 PPPRAPPRA2P REEP PEPP	WA MADDALLERESSEED 5 HHOREELEENO DODA XSSS 741	ET FEFFERENMENNENNENNENNENNEN RARK TTT	FERENCIEVE I IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	THE PARTY PATTER A THREE THE TAR FREE S THE	PULIVOLDELEDIN L IIIIIIIII LILL III III 2 233	UPDEPERTURE LEALENEEDEL ITTV ATT 245	WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW	247	5 T	5 249	EFFERERERERERERERERERERERERERERERERERER	AUGGEGEBEEEEE EEEEEEEEEEEEEEEEEEEEEEEEEEE	CC CFCCCCCSDCCCS S SSSSSSSSSS FEET TOOL 2 251	KK AMAAMAAASSAA P PROPAGASSAS 2275 CCCC 252	SSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS	V HHHH 0000 00000000 N HHHH V	LC, I THILLIII I IIIIIIIII I MITHATTATATA	255 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	KRRRRRRRRRRR A CONCOMMENT A 256	LITITITITITIE CONCERNMENT UNIO EEEE A 257	AVYVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVV	PROPERTY AND	11 UNINGUEGOCOG 6 GEOGGEOGGEOG GAZE 1714 259	EFFEFEFEFEFEFEFEFEFEFEFEFEFEFEFEFEFEFE	YTYYYYYYYYYYY U TO THE FEFT A TO THE FEFT A TO THE A	WI CONSTRUCTION & XXXXXXXXXX KKKK		ANDERDERERE RECERERERE FOR FOR FOR THE TANK 763	ICCCCCCCCCCCCC + renewand and back back A 264	NERREARE COLOCIC CONC CCCC A COLOCICIÓN ANALANA	10 202 THE REAL PRESERVER A PRESERVER AND A PRESERVER A PRESERVERA	TH RECHARGERER N RECERCERERERERERERERERERERERERERERERERE	XX RERERERERER X KKEYEVEN XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX		, T YTYYYYYYY YYYY YYYY . 268 J	269		210 200 2000 2000 2000 2000 210 210 210	THE STATE AND A SAMANANAN SAMA SAMA SAMA SAMA SAMA SA	IIIIIII NUVVVIS SAAAAAAAA	REARENEWERE A 212 3	273 A PROVININAN TTET REAL A 273	WWWWWWW V WWWWWWW	THE REPORT OF THE PARTY OF THE	ED HKKKKKKKKKKKKK O KKKKKKKKK SODU A	KA NTTTTTTTIIOTH Y MAYYYYYYYYY RARK DEEE A	FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF	THEFTER PROPERTY AND A REPORT OF COLOR	ER DDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD	CONTRACTOR DECE A	VULLEVELTTY & ELECTERELE IVV AAAA 201		NW NNNNNNNNNNNNNNN DDDDDNHNNNN NAA TATTTY A 542	MOCCCCCCCCTT Y YYYYYYYYY HUDD 10000 281	284 284 287 284	an anarcherthiaste L FFLEItrist		Addd Leer Thirty I TAWMAIIAIIMA	ARARARSARANAN C TURNELLE LELL B 407	ANAMAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	LUTITITITIE AND AND AND AND ANA AAAA B 289	VIIIIIIII I LULIULULULULULULULULULUU		AN GUDDDDDDDDEEAO D DDDDDDDDD 211 MLUL B	GG EEEEEEEDNYTCH N WY WILL EEEEEEDNY	DO BEAUER CONTRACTOR DI MISCOCOCOCO CALUE 0000 241	UN THAT AND		FEFFERLLLEF & PEPPERSTER ALL FYFE 9 295	CUCTCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	CHEMICATION C LUCCOCCCC CCCC B 297	WINGSTRUCTURE OF A LIGHTLEFELL WAVY VAVV R 298	ОСССООЛОНИИНИИНИИ И ИМИНИИИ И ИМИНИИ И И И И И И	6 6666666757 Cree cree	<u>6.0000000000 0000</u>	1111111111111111	22222222222222222222222222222222222222	TT TREPREPEREPEREPEREPEREPEREPEREPEREPEREPE	SECONDOCCERENT N HASSSSSSSS SSDE	
	CC TWFTTTTTTTTTT TTTTTTTTTT TTST CCCC A	Y LLELELITITITI STTTTTTS LLET LLLE A 230	LEGENERALELELELE VYLERELELELE LELE YYYY A 231	Y FFLULLLLLLLLLLL I VVVVVVVV LILLM WWWW A 212	YYYYYYYYYYYY AAGAAAAAAA AACC AVVV A 234	NK KNNKNKKKKKKK K KKKKKKK I KKKKKKKK I KKKKKKKK	The VIIIIIIVIILLUI L'UNUNUNUTI IIIVII	NIN KKKKKKKKKKSSKK R RRRURARARARA BASK LT. 237	BEZ A YAYYYYYYYYYYYYYYYYYYYYYYYYYYYYY AN DA 2	DE LETTERATUREPERK 2 PPPRAPPRA2P REEP PEPP	W WARDELEESEED S HHOEEEEENO DDA XSSS 341	ET FEFFERENMENNENNENNENNENNEN RARK TTT	FERENCIEVE I IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII		PULIVOLDELEDIN L ILTITITITI LELL THE CONTRACT	UPDEPERTURE LEALENEEDEL ITTV ATT 245	WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW	247	5 T	5 249	EFFERERERERERERERERERERERERERERERERERER		CC CFCCCCCSDCCCS S SSSSSSSSSS FEET TOOL 2 251	KK AAVAAAAAASAA BERCUCUSUUS SUTS CCCC 252	0.01 HH SSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS	TI TITE A COOL COOLOGOOD DAVE CONTRACT TITE	LC, I THILITIIII I MITTATION	255 222 222 222 222 222 222 222 222 222	0.46 SS KRERRERERER O CONCOUNT OF 1 1 256	Y IIIIIIIIIII S CONTRACTOR UNK EEEE A 257	AVYVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVV	D. D. T. TITTTTTTTTTTTTTTTTTTTTTTTTTTTTT	11 UNINGUEGOCOG 6 GEOGGEOGGEOG GAZE 1714 259		YTYYYYYYYYYYY U TO THE FEFT A TO THE FEFT A TO THE A	THE PROPERTY A TYPYTYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY		ANDERDERERE RECERCERERE FOR FOR FOR 263	WH ICCCCCCCCCC & TOTOTOTOTOTOTOTOTOTOTOTOTO	NERREARE COLOCIC CONC CCCC A COLOCICIÓN ANALANA	D 14 UL VILLE AND	PURKHURRARRARRA N RERARRARRAR DEEN III 266	U-21 KK PRRARARARARARARARARARARARARARARARARARAR		, T YTYYYYYYY YYYY YYYY . 268 J	269		210 BOOK BOOK BOOK BOOK 210	THE STATE AND A SAMANANAN SAMA SAMA SAMA SAMA SAMA SA	I A A A A A A A A A A A A A A A A A A A	DE REMERANDATION AND THE SECOND AND	273 A PROVININAN TTET REAL A 273	WWWWWWW V WWWWWWW WWW	A A A A A A A A A A A A A A A A A A A	HKKKKKKKKKKKK O KKKKKKKKKE SOUL	0.42 KA NTTITTIIOTH Y MAYYYYYYYYYYY RARK DEEE A	CC FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF	THEFTER PROPERTY AND A REFERENCE COCC	DDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD	SE STATESTATESTATESTATESTATESTATESTATEST	20 CONCOLOCITIV & ELECTICIC IIVY AAAA 281		VI NW WWWWWWWWWWW D DDDDNDDDDDDDDDDDDDDDD	NV MCCCCCCCCTT Y YYYYYYYYY HOUD DODD	The second	a subbeletingenet L FFLETISTER		ddd tere and a firm a manual alling an	. VB AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	43 AA ANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	LUTITITITIE AND AND AND AND ANA AAAA B 289	VIIIIIII I LULIDULULULULULULULULULULU		CODDDDDDDDEEAO D NDODDDDDDD 2011 NLUL B	GG EEEEEEEDNYTCH N WY WILL EEEEEEDNY	DO BEAUER CONTRACTOR DI MISCOCOCOCO CALUE 0000 241	UN THAT AND	The second secon	Le Preterezhilter & Freesennen and Free B 295	CC CCTCTCCCCCCCCCC CCCCCCCCCCCCCCCCCCC	W CLANNER CONCOUNTER CONCOUNTER 297	WINGSTRUCTURE OF A LIGHTLEFELL WAVY VAVV R 298	ОСССООЛОНИИНИИНИИ И ИМИНИИИ И ИМИНИИ И И И И И И	6 6666666757 Cree cree	AG CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	22222222222222222222222222222222222222	PEPEPEPEPEPEPEPEPEPEPEPEPEPEPEP	SECONDOCCERENT N HASSESSES SEDE	
	CC TWFTTTTTTTTTT TTTTTTTTTT TTST CCCC A	Y LLELELITITITI STTTTTTS LLET LLLE A 230	The restocateductor a statements for yyyy A 231	Y FFLULLLLLLLLLLL I VVVVVVVV LILLM WWWW A 212	60 LL YYYYYYYYYYYYYY AAGAAAAAAAA AACC AVVV A 234	NK KNNKNKKKKKKK K KKKKKKK I KKKKKKKK I KKKKKKKK	The visitivities in the second	NIN KKKKKKKKKSSKK R REGURARAR RURK III A 237	BEZ A YAYYYYYYYYYYYYYYYYYYYYYYYYYYYYY AN DA 2	DE LETTERATUREPERK 2 PPPRAPPRA2P REEP PEPP	W WARDELEESEED S HHOEEEEENO DDA XSSS 341	ET FEFFERENMENNENNENNENNENNEN RARK TTT	FERENCIEVE I IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	The second secon	PULIVOLDELEDIN L ILTITITITI LELL THE CONTRACT	245 2 242 2 242 2 242 2 242 2 242 2 242 2 242 2 245	WA NAVANAPRARARA R RARARARA SOCIETATION 246	the choletononecon d preserve many many 3 47	AIA HINDRINGNAMAN A A A A A A A A A A A A A A A A A	5 249	Ста станатичны и нининини и ининининини станата станатичны и и и и и и и и и и и и и и и и и и и		CC CLCCCCCSDCCCS S SSSSSSSSSS EET 2000 2 231	KK AAVAAAAAASAA BERCUCUCUCUCUCUCUCUCUCUCUCUCUCUCUCUCUCUCU	0.01 HH SSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS	TI TITE A COOL COOLOGOOD DAVE CONTRACT TITE		255 222 222 222 222 222 222 222 222 222	0.46 SS KRERRERERER O CONCOUNT OF 1 1 256	Y IIIIIIIIIII S SUCCESSION UNK EEEE A 257	ALA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	D. D. T. TITTTTTTTTTTTTTTTTTTTTTTTTTTTTT	11 MARCOGGGGGG G GGGGGGGGGG GALL 11 259		KK YYYYYYYYYYYYYY C C C C C C C C C C C	WI CONSTRUCTION & XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX		CE PAREEDEEDEELE E EELEREFEE FEF FOR 263	WH ICCCCCCCCCC & TOTOTOTOTOTOTOTOTOTOTOTOTO	TI REFERENCESS CONCESSION A 11	D 14 UL VILLE AND	TH RECHARGERER N RECERCERERERERERERERERERERERERERERERERE	U-21 KK PRERERERERERERERERERERERERERERERERERERE	YT HYYYYYYEESSUUL A ANALANANAKK KWKW KKKW A	, T YTYYYYYYY YYYY YYYY . 268 J	269		210 200 2000 2000 2000 2000 210 210 210	110 SANA SANA SANA SANA SANANANANAN SANA S	I A A A A A A A A A A A A A A A A A A A	DE REMERANDATION AND THE SECOND AND	273 X TAT THE TATE WANNING A NUMBER AND THE REAL X 273	WWWWWWW V WWWWWWW WWW	A A A A A A A A A A A A A A A A A A A	ED HKKKKKKKKKKKKK O KKKKKKKKK SODU A	0.42 KA NTTITTIIOTH Y MAYYYYYYYYYYY RARK DEEE A	CC FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF	CC THTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	ER DDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD	CONTRACTOR DECE A	VULLEVELTTY & ELECTERELE I IVY AAAA 201		1.21 NW WWWWWWWWWWW D DDDDDDDDDDDDDDDDDDDDD	NV MCCCCCCCCTT Y YYYYYYYYY HOUD DODD	284 284 287 284	a subbeletingenet L FFLETISTER		PALITVITVITVIT I UTIVITVITVITVITVITVITVITVITVITVITVITVITVI	. VB AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	AN ANAMAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	LL LVTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	VIIIIIII I LULIDULULULULULULULULULULU		AN GODDDDDDDDEEAO D DDDDDDDDD 211 MLUL B	GG EEEEEEEDNYTCH N WY WILL EEEEEEDNY	DU PERARATATION DI MISCOCOCODI CUMO 0000 241	W PROMARKARKK E KKEODODAR RESEARCH	The second secon	Le Preterezhilter & Freesennen and Free B 295	CUCTCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	2.37 VV CHANNELING COLOCOCCCC COCC B. 297	HU MARGING AND AND AND AND AND AND A 298	A C C C C C C C C C C C C C C C C C C C	HALL TAXAGAGAGAGAGA G GOOGOOGOTATO DITES 233	2 KA	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	2222 2522252525 2 25252525252525255 27	TT TREPREPEREPEREPEREPEREPEREPEREPEREPEREPE	SECONDOCCERENT N HASSESSES SEDE	
	CC TWFTTTTTTTTTT TTTTTTTTTT TTST CCCC A	2.60 YY LILLELELELETTING STTTTTTSS LLET LELE A 229	The restocateductor a statements for yyyy A 231	0.40 CALLELELELELE I VVVVVVVI LILM WWWW A 212	60 LL YYYYYYYYYYYYYY AAGAAAAAAAA AACC AVVV A 234	K XX KNKKKKKKKKKKK K KXXXXXXXXXXXXXXXXX	The visitivities in the second	NIN KKKKKKKKKSSKK R REGURARAR RURK III A 237	5. 40 M TATATATATATATATATATATATATATATATATATAT	1.12 TO TOTAL TOTAL TOTAL TOTAL TOTAL TOTAL TOTAL	0.67 UN DECEMBERED S HIDEEREENO DODA X555 741	2 60 TI WARMANNANNAN TIWA RARARARARARA RARA TIT 743	1.92 Generative I IIIIIIIII III LILL P 202	The second secon	PULIVOLDELEDIN L ILTITITITI LELL THE CONTRACT	245 2 242 2 242 2 242 2 242 2 242 2 242 2 242 2 245	WA NAVANAPRARARA R RARARARA SOCIETATION 246	the choletononecon d preserve many many 3 47	AN HINDRAWANANAN A ANALANANANANAN A ANALANANANANAN	S RUNN UKUN KUNNYANANA ATATATATATATATATATATATATATATATATAT	Ста станати и нинининини и инининитери		CC CLCCCCCSDCCCS S SSSSSSSSSS EET 2000 2 231	V.U. KK AMAMAMAASSAA PERECONSUS 227 CCC 252	0.01 RH SSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS	2.60 LI TITTITI UQQ000000 0000 D 11 11 11 11 11 10 00000000		255 222 222 222 222 222 222 222 222 222	0.46 S5 KRARARARARARARA O CONCOUNT 111 A 256	VY INTITITITIES SUCCESSION UNK EEEE A 257	2.60 FF VYVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVV	D. D. T. TITTTTTTTTTTTTTTTTTTTTTTTTTTTTT	11 MARCOGGGGGG G GGGGGGGGGG GALL 11 259		KK YYYYYYYYYYYYYY C C C C C C C C C C C	WI CONSTRUCTION & XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX		CE PAREEDEEDEELE E EELEREFEE FEF FOR 263	WH ICCCCCCCCCC & TOTOTOTOTOTOTOTOTOTOTOTOTO	TI REFERENCESS CONCESSION A 11	D 14 UL VILLE AND	TH RECHARGERER N RECERTERERERERERERERERERERERERERERERERER	U-21 KK PRERERERERERERERERERERERERERERERERERERE	1.90 YT HYYYYYEFELLON THANARAK KWKW KKKW A	, T YTYYYYYYY YYYY YYYY . 268 J	269			110 SANA SANA SANA SANA SANANANANAN SANA S	I A A A A A A A A A A A A A A A A A A A	0.62 DE REKERKERKERKERKER	273 X TAT THE TATE WANNING A NUMBER AND THE REAL X 273	WWWWWWW V WWWWWWW WWW	A A A A A A A A A A A A A A A A A A A	C 10 HERKERKERKERKERKERKERKERKERKER	0.42 KA HITTHITTIIOTH Y MANAGARANA MANA DEEE A	2.37 CC FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF	1.17 CC TITTETTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	0.70 ER DUDDODODODODOD I IIIIIIII IICC	0.36 Se	2.60 2. 2. 2. 2. 2. 2. 2. 2. 2. 2. 2. 2. 2.		12 A TIT TIT CONTROLLING DEPENDING NUMBER NO 12-0	AT MCCCCCCCC A AAAAAAAAAAAAAAAAAAAAAAAAA	The second	a subbeletingenet L FFLETISTER		ddd fere and the first internation of the first search and the first sea	AA AAAAAAASAAAAAA AAAAAAAAAAAAAAAAAAAA	43 AA ANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	LL LVTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	2.60 New VIII I LUCIDECULUE III LUCIDECULUE ZIJI LULI E 290	a start iti ilaanaana i laalilaaliila ta		0.17 CALEREFERINGEN CONTRACTOR DURINGER DUND NNNN	0.81 00 BERNARD CONSTRUCTION OF BUILDING CONSTRUCTURE OF BUILINA CONSTRUCTURE OF BUILONO OF BUILO OF BUILO CONSTRUCTURE O	2 48 WILLIAM CONTRACTOR RECEIPTING		LE REFERENCIAL & FRENCH 2 200 200 200 200 200 200 200 200 200	CC CCTCTCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	2.33 VV CHARACTERISTIC COCCCCCCCCCCCCCCC B. 297	H LINE AND A LINE LEVEL AND A LINE LEVEL AND AND A 298		ALL MASKAGAGAGAGA G GGGGGGGGGGGGGGGGGGGGGGGG	2 KA	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	2222 2522252525 2 25252525252525255 27	TT TREPREPEREPEREPEREPEREPEREPEREPEREPEREPE	SECONDOCCERENT N HASSESSES SEDE	
016 1.10 1.10	2.60 0.33 CC TWVFITFITTITT T TTTTTTTTTT TST CCCC A	2.60 YY LILLELELELETTING STTTTTTSS LLET LELE A 229	143 I 2.60 th reducedenced a state 121 Yay A 231	144 0.40 Y FFLELLELELELEL I VVVVVVVVI LILEM MWWW A 212	2.60 LL YYYYYYYYYYUU V ANGANAAAA AACC AVVV A 234	145 K X KNEWKRENKER & WELEVELETELETER A 235	2 5 14 2 2 5 14 2 2 5 14 2 2 5 14 2 2 5 14 2 2 5 14 2 2 5 14 2 2 5 14 2 2 5 14 2 2 5 14 2 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 5	149 0.67 NN NNKKWKWKKSSKK R RURURSRAM ROW 111 A 237	150 c concernant of the concernation of the concernati	151 1.12 DE LECENTREPERT E PERTEPERTE E PER PERE	152 0.67 du municipalete Seeto 5 HHOREELEENO DDA X555 741	153 Z 60	154 1.92 Generativerse International International Second	155 2.60	156 2 60 WEILINGLEEVENING LITTITITITITITIE 244	157 R 245 245 245 245 245 245 245 245 245 245	C C C C C C C C C C C C C C C C C C C	tet doddogogogo d propresson man 247	AN HINDRAMANANANAN I TATATATATATATATATATATATATATATATATATAT	A A A A A A A A A A A A A A A A A A A	тория и при нарании и при при при при при при при при при		A A CULCCCCCSDCCCS S SSSSSSSSSSSSSSSSSSSSSSSSS	KK AMAMANANSSAN DEPENDENT	0.01 RH SSSSSSSSTETER AND ADDRESS RAPE 253	103 2.60 LI TITTITUTUTUTUTUTUTUTUTUTUTUTUTUTUTUTUT	166 0.16 D. THILITIN I TILITINI I TILITIN C. 201	255 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	0.46 S5 KRARARARARARARA O CONCOUNT IIII A 256	YY IIIIIIIIIIIII & CONSTRUCTION UNKO EEEE A 257		D D D D D D D D D D D D D D D D D D D	1 2 60 11 10 0000000000 0 00000 0000 11 259		KK YYYYYYYYYYYYY CEREFEER A 261	WI CONSTRUCTION & XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX		CE PAREEDEEDEELE E EELEREFEE FEF FOR 263	WH ICCCCCCCCCC & TOTOTOTOTOTOTOTOTOTOTOTOTO	TI REFERENCESS CONCESSION A 11	D 14 UL VILLE AND	TH RECHARGERER N RECERCERERERERERERERERERERERERERERERERE	267 T XX REARARARARARARARARARARARARARARARARARARA	1.90 YT HYYYYYEFELLON THANARAK KWKW KKKW A	, T YTYYYYYYY YYYY YYYY . 268 J	269		210 200 2000 2000 2000 2000 210 210 210	1 C SANARANANANA SANARANANANANANANANANANANANANANANANANANA	I A A A A A A A A A A A A A A A A A A A	0.62 DE REKERKERKERKERKER	2.60 * **********************************	WWWWWWW V WWWWWWW WWW	THE PROPERTY AND THE PR	C 10 HERKERKERKERKERKERKERKERKERKERKERKERKERKE	0.42 KA HITTHITTIIOTH Y MANAGARANA MANA DEEE A	C FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF	1.17 CC TITTETTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	ER DDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD	0.36 Se	I 2.60 CONCENTERING LARGELEILLELL IVY AND 201		1.21 NW WWWWWWWWWWW D DDDDDDDDDDDDDDDDDDDDD	AT MCCCCCCCC A AAAAAAAAAAAAAAAAAAAAAAAAA	2.60 2.60 2.64 2.84	a subbeletingenet L FFLETISTER		ddd tere and a firm a manual alling an	AA AAAAAAASAAAAAA AAAAAAAAAAAAAAAAAAAA	0.43 AN ANALALATION CONTRACTOR SCG ANAL B 288	LL LVTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	2.60 BM UTTERVENTION I LULULULULUL III LIA B 290		AN GODDDDDDDDEEAO D DDDDDDDDD 211 MLUL B	0.17 CALEREFERINGEN CONTRACTOR DURINGER DUND NNNN	0.84 On Berner Contraction Division Range 0000 291	W PROMARKARKK E KKEODODAR RESEARCH	The second secon	LE REFERENCIAL & FRENCH 2 200 200 200 200 200 200 200 200 200	CC CCTCTCCCCCCCCCC CCCCCCCCCCCCCCCCCCC	2.33 VV CHARACTERISTIC COCCCCCCCCCCCCCCC B. 297	H LINE AND A LINE LEVEN A LINE LEVEL AVAY AVAY A 298	A C C C C C C C C C C C C C C C C C C C	HALL TAXAGAGAGAGAGA G GOOGOOGOTATO DITES 233	2 KA		2222 2522252525 2 25252525252525255 27	TT TREPREPEREPEREPEREPEREPEREPEREPEREPEREPE	SECONDOCCERENT N HASSESSES SEDE	

#### T.F. JENNY ET AL.

TABLE I. Secondary Structure Prediction for the Serine/Threonine Specific Protein Phosphatases

Segment	Structure	Comments
003-019	Helix	Reliable, perfect amphiphilicity
021-022	Parse	PN, PS, string of 5 surface assignments
023-026	Coil	In some protein family members, possibly an edge strand
027-029	Parse	SPP, DSP, PXP, NP, SP, conserved P
030-034	Beta	Near active site region 035–042
035-042	Active site	Conserved D, HGQ, D
043-049	Alpha	Near active site region 035-042
050-058	Parse	Gap, PPSSN, GGDP, GGSP
059-062	Strand	Reliable
063-074	Active site	
075-086	Helix	Internal helix
087-088	Parse	PS, PN, PD
090-093	Strand	Reliable, central in a sheet
094-098	Active site	Conserved RGNHE
101–116	Helix	Reliable, residue 107 assignment must lie on the surface
117-120	Parse	Gap, GGNS, GNS
19-130	Helix	Reliable
134	Parse	Weak, partly conserved P, break in amphiphilicity of preceding heli
135-139	Strand	Reliable
40-142	Parse	NNS, DG
143–146	Strand	Reliable
147-152	Active site	Conserved HGGXSP
151-153	Parse	SPS, SPD
154-166	Helix	310 geometry suggested by patterns of conservation and variation
169-174	Parse	PDSG, GGP, PP, GP
175-180	Strand	Reliable
181-225	Parse	SDPSGD, NNNP, PG, SP; surface loop distant from active site
228-232	Strand	Reliable, core of β-sheet
233-235	Parse	GPD, GSD
236-243	Helix	Reliable, surface
243-245	Parse	NNG, NŃN
246-262	Helix	Middle passes near active site
263-273	Parse	Gap, DGG, PS
274-278	Strand	Unusual patterns of conserved functional groups
279-283	Active site	Conserved SAPNY
284-287	Coil	
288-289	Parse	GN .
290-295	Strand	
297 <b>-299</b>	Parse	DDS, end of core alignment

ends at the active site. These are therefore assigned as a supersecondary structural unit formed by two antiparallel structural units, a helix and a strand. The most plausible structural model assembles this supersecondary structural unit with the preceding supersecondary structural unit to yield a parallel  $\beta$ -sheet. These assignments suggest that this part of protein phosphatase forms an  $\alpha$ - $\beta$  parallel sheet fold, well known in structural biology. Assignments of the remaining supersecondary structural elements are less reliable, and await confirmation by a detailed covariation analysis.

One application of secondary structural models is to detect distant homology between protein families that is not established by statistically significant sequence similarities. In this application, the predicted secondary structure from one family of proteins is aligned either with a predicted secondary structure or an experimental structure from another protein family that shares poor sequence similarity. For example, in its search of the database, DARWIN identified and aligned six positions (169–174) of the serine-threonine protein phosphatases with a tyrosine-specific protein phosphatase (a PPSHAP sequence), suggesting that the tyrosine and the serine-threonine protein phosphatase families might be homologous despite their apparent mechanistic differences. Superimposition of the secondary structural elements of the two protein families, one predicted and the other experimental, clearly shows that these two families are not homologous, however. This suggests that this particular sequence similarity in this one tyrosine phosphatase arose by convergent, not divergent, evolution.

Finally, several distinct mechanistic classes of enzymes are known that catalyze the transfer of a

### SERINE/THREONINE-SPECIFIC PROTEIN PHOSPHATASES













- Cohen, P. The structure and regulation of protein phosphatases: Annu. Rev. Biochem. 58:435-508, 1989.
   Barford, D., Keller, J. C. Cocrystallization of the catalytic subunit of the serine/threonine specific protein phosphatase 1 from human in complex with microcystin LR. J. Mol. Biol. 235:763-766, 1994.
   Gonnet, G. H., Benner, S. A. Computational biochemistry research at ETH. Technical Report 154, Departement Informatik, March, 1991.

- research at ETH. Technical Report 154, Departement Informatik, March, 1991.
  29. Gerloff, D. L., Jenny, T. F., Knecht, L. J., Gonnet, G. H., Benner, S. A. The Nitrogenase MoFe Protein: A secondary structure prediction. FEBS Lett. 318:118-124, 1993.
  30. Jenny, T. F., Benner, S. A. Evaluating predictions of secondary structure in proteins. Biochem. Biophys. Res. Commun. 200:149-155, 1994.
- Pavone, V., Di Blasio, B., Santini, A., Benedetti, E., Pedone, C., Toniolo, C., Crisma, M. The longest, regular polypeptide 310 helix at atomic resolution. J. Mol. Biol. 214: 633-635, 1990.
   Barlow, D. J., Thornton, J. M. Helix geometry in proteins. J. Mol. Biol. 201:601-619, 1988.
   Kim, E., Wyckoff, H. W. Reaction mechanism of alkaline phosphatase based on crystal structures. J. Mol. Biol. 218: 449-464, 1991.

- phosphatase based on crystal structures. J. Mol. Biol. 210.
  449-464, 1991.
  34. Benner, S. A., Glasfeld, A., Piccirilli, J. A. Stereospecificity in enzymology: Its place in evolution. Topics Stereochem. 19:127-207, 1989.
  35. Russell, R. B., Barton, G. J. The limits of protein secondary structure prediction accuracy from multiple sequence alignment. J. Mol. Biol. 234:951-957, 1993.