ELSEVIER

Biochimica et Biophysica Acta 1314 (1996) 191–225

BIOCHIMICA ET BIOPHYSICA ACTA
BBA

Review

How do protein kinases recognize their substrates?

Lorenzo A. Pinna *, Maria Ruzzene

Dipartimento di Chimica Biologica, Università di Padova, and Centro per le Biomembrane del CNR, Padua, Italy

Received 10 April 1996; revised 12 June 1996; accepted 13 June 1996

Contents

Abbreviations: AMP-K, AMP-activated protein kinase; βARK, β adrenergic receptor kinase; BCKDHK, branched chain ketoglutaric acid dehydrogenase; CaM-kinases, calmodulin-dependent protein kinase; cdc2 kinase, protein kinase expressed by CDC2 gene; cdk, cyclin dependent kinase; CK1, protein kinase CK1 (also termed casein kinase 1 or I); CK2, protein kinase CK2 (also termed casein kinase 2 or II); CSK, C-terminal Src protein kinase; dSRNA-PK = PKR; EGF, epidermal growth factor; ERK, extracellular signal-regulated kinase; FGF, fibroblast growth factor; GSK3, glycogen synthase kinase-3; HCR, heme controlled repressor (phosphorylates eIF2α); HMG-CoA Red. kin. A, HMG-CoA reductase kinase A; HRI = HCR; HSV, herpes simplex virus; IGF1, insulin-like growth factor-1; I-R, insulin receptor; MAP kinase = ERK; MAPKAP-kin.1 = $p90^{s6k}$; MAPKAP-kin.2, MAP kinase activated protein kinase 2; MBP, myelin basic protein; MLCK, myosin light chain kinase; Nek, *N*IMA-r*e*lated *k*inase; NIMA, *n*ever *i*n *m*itosis protein kinase; $p70^{s6k}$ and $p90^{s6k}$, 70 and 90 kDa kinases that phosphorylate s6 protein; PDGF, platelet derived growth factor; PDHK, pyruvate dehydrogenase kinase; PhK, phosphorylase kinase; PK, protein kinase; PKA, cAMP-dependent protein kinase; PKB = RAC; PKG, cGMP-dependent protein kinase; PKI, heat stable inhibitor protein of PKA; PKR, RNA-dependent protein kinase; PRK1, protein kinase C-related kinase 1; PTK, protein tyrosine knase; RAC, PK related to PK*A* and PK*C*; RhK, rhodopsin kinase; SNF-1 PK, sucrose non-fermenting protein kinase; TGF, transforming growth factor

* Corresponding author. Dipartimento di Chimica Biologica, Via Trieste, 75, 35121 Padua, Italy. Fax: 39-49-8073310; e-mail: pinna@civ.bio.unipd.it.

PII S0167-4889(96)00083-3

1. Premise

Three excellent reviews on the recognition of sequence motifs by protein kinases were published in 1990 [1] and 1991 [1a,2], a circumstance that may call into question the timeliness and utility of a new review covering more or less the same ground. It should be borne in mind, however, that in the few years that have since elapsed the number of known protein kinases has more than doubled and the first crystal structures of these enzymes have been solved, disclosing the structural features that may underlie their substrate specificity. At the same time new insights have been provided into the mechanism by which protein-protein interaction modules such as Src homology domains-2 and -3 (SH2 and SH3) [3], pleckstrin homology (PH) domains [4,5] and phosphotyrosine binding domains (PID or PTB) [6,7] may contribute to the recruitment of phosphoacceptor substrates, and revolutionary methods for the study of protein kinases site specificity based on the use of peptide libraries have been established [8–15].

The aims of this paper therefore are, on the one hand, to present and update the information provided by Kemp and Pearson [1] and by Kennelly and Krebs [2] on the specificity of additional protein kinases and the development of new tools and concepts; and on the other hand to discuss the structural features of protein kinases that are now believed to be implicated in substrate recognition. At the same time, some basic concepts underlying substrate recognition by protein kinases, such as site specificity, consensus sequence, substrate recruitment, sequential phosphorylation, and their relevance to physiological situations, will be critically examined.

2. Site specificity: a general feature of protein kinases?

Protein kinases are, with occasional exceptions, members of a huge superfamily of enzymes sharing with protein phosphatases the responsibility of regulating in a concerted manner virtually every kind of cellular function. Having proteins as substrates and having to affect only one or few of the many potential phosphoacceptor residues in their targets, protein kinases must be endowed with a tremendous selectivity, even higher than that of protein phosphatases which, after all, are 'only' required to make their choice from a smaller variety of phosphorylated targets, pre-selected by the kinases. More than 300 protein kinases are known to date (February 1996) and it has been predicted that their final number will be between 2000 and 3000 [16]. Even so they are less abundant than their substrates, since 30 to 50% of the proteins of a eukaryotic cell undergo phosphorylation, as judged by in situ radiolabeling followed by

2D electrophoresis. These figures would mean that, on the average, each kinase should have a dozen or so substrates, and indeed all available evidence shows that protein kinases, with few exceptions, are in general pleiotropic enzymes impinging on a wide spectrum of protein substrates with varying structures and often implicated in different functions.

Such a situation supports the idea that the majority of protein kinases are able to recognize a number of structural features surrounding the target aminoacids in all their protein substrates, which would imply that the integrity of the substrate is not required since recognition is based on a common local motif shared by all the proteins targeted by a given kinase. The validity of this concept was proved with a wide variety of protein kinases by showing that relatively small peptides reproducing the phosphoacceptor sites can be phosphorylated with kinetics comparable to those of the intact protein substrate.

Site specificity is not the only tool ensuring the selectivity of protein kinases, however. Subcellular compartmentation and specific association mediated by targeting elements outside the catalytic domain have been shown to play a crucial role in increasing further the selectivity of protein kinases, and even more of protein phosphatases [17]. The example of PKA, endowed with a marked site specificity and with a tight mechanism of regulation, whose subcellular targeting is nevertheless controlled by a family of A-kinase anchoring proteins (AKAPs) is particularly illuminating in this respect [18]. Supramolecular targeting devices, however, are outside the scope of this review, predominantly devoted to those features that determine or in any way affect the *direct* recognition of protein substrates by kinases. It should be anticipated that in some instances the recognition elements determining the interaction with a protein kinase are not entirely located in the primary structure of the phosphoacceptor site, but either constitute a separate recognition module or belong to the phosphoacceptor site by virtue of the folded tertiary structure of the protein substrate. In these cases, which will be discussed below (Section 8), the peptides reproducing the primary structure around the phosphorylated residue will prove poor substrates compared to the intact protein.

It is possible that similar allosteric and/or three-dimensional recognition motifs as yet unrecognized, account for the extreme specificity of a few protein kinases that are almost inactive, not only towards peptides reproducing their phosphoacceptor sites, but even towards the denatured protein substrates.

If the ability to phosphorylate peptides reproducing phosphoacceptor sites is held as the overriding criterion of site specificity, then it can be concluded that the large majority of protein kinases do display 'site specificity'. Both the relative efficiency of peptide phosphorylation compared to the intact protein, and the relevance of local structural features are quite variable, however, among different protein kinases. Thus, if more stringent criteria are adopted, requiring phosphorylation efficiency comparable to that of natural substrates and 'indispensability' of given specificity determinants, the number of protein kinases endowed with marked 'site specificity' decreases, being unevenly distributed between Ser/Thr kinases, the majority of which are markedly site specific, and Tyr kinases, where a well-defined site specificity is rather the exception.

2.1. Residue and sequence specificities

Efficient phosphorylation of aminoacids by protein kinases in general requires their incorporation into a discrete peptide sequence or, rarely, in a larger polypeptide chain. During the catalytic event the residue undergoing phosphorylation will bind to the active site, whereas the surrounding residues will be accommodated nearby in a substrate binding groove, interacting with elements of the kinase outside the sensu stricto catalytic site of the enzyme. Both the nature of the phosphorylatable aminoacid, therefore, and that of the residues surrounding it will influence the catalytic event, giving rise to 'residue specificity' and 'sequence specificity' respectively, the former being also referred to as 'active site specificity' [19]. The overall 'site specificity' will reflect both 'residue/active site' and 'sequence' specificities.

2.2. Ser / Thr and Tyr specific protein kinases

Despite their common architecture, the active site conformation of protein kinases may vary quite significantly, as outlined on the one hand by their strikingly different sensitivity to the competitive inhibitor, staurosporine [20], and on the other hand by elegant studies with alcohol-bearing compounds that

can serve as protein kinase substrates [19,21,22] and are proving useful to explore the active site specificity of individual kinases [23,24].

From a practical standpoint, however, the active site specificity, subtle as it might be, gives rise in nature to just two types of residue selections corresponding to the two large classes of protein kinases, either Ser/Thr or Tyr specific. These include nearly all members of the huge eukaryotic protein kinase family [25] as well as few unorthodox protein kinases functionally related to the Ser/Thr specific group but structurally similar to so called 'histidine kinases' of prokaryotes [26].

In most instances, the borderline between Ser/Thr and Tyr specific protein kinases, empirically drawn using substrates of the two sorts, is quite sharp and reliable. It is also based on differences in primary structure motifs [27] whose rationale is being unravelled as the crystal structure of a protein tyrosine kinase, the catalytic domain of the insulin receptor, has been solved [28] and can be compared to that of the paradigm Ser/Thr kinase, the catalytic subunit of PKA [29].

In contrast, the features that underlie the preference for either seryl or threonyl residues within the class of Ser/Thr specific protein kinases remain unclear. As a general rule Ser/Thr protein kinases tend to prefer seryl over threonyl residues, in some instances to such an extent that it would be tempting to postulate the existence of a sub-class of just Ser-specific kinases, virtually inactive on threonine. These are, e.g. the Golgi apparatus casein kinase, GEF-CK [30] and $p90^{s6k}$ (= MAPKAP kinase 1) [31]. On the other hand, other kinases, e.g. AMPK [32] and cdc2 [33] tolerate quite well Ser → Thr substitution and PYT/ESK seems to phosphorylate proteins preferentially on Thr [33a].

It should finally be mentioned that besides the protein kinases that catalyze the phosphorylation of the hydroxyl residues of Ser, Thr and Tyr, protein kinases specific for Arg [33b] and for His, which may be not related to the so-called histidine kinases of the two component system of prokaryotes [26], have also been described in eukaryotes (e.g. [34]). These enzymes, which may play an important role in novel signal transduction pathways [35] are still awaiting structural characterization and will not be considered here.

2.3. Dual specificity protein kinases: much ado about nothing?

The discovery that the activity of cyclin dependent kinases and MAP kinases is tightly controlled, though in opposite direction, by the simultaneous phosphorylation of a threonyl and a tyrosyl residue adjacent or very close to each other led to the detection some years ago of so-called 'dual specificity' protein kinases [36] capable, in principle, of phosphorylating either Ser/Thr or Tyr residues. An appealing hypothesis connected with that discovery was that the signal transduction pathways, initiated by tyrosine phosphorylation at the level of receptor protein tyrosine kinases and ending with a dissemination of Ser/Thr phosphorylation within the cell, had to pass through an intermediate step mediated by dual specificity protein kinases. While this hypothesis was soon dismissed after the finding that in the MAP kinase signalling cascade the dual specificity MAP kinase kinase (MEK) is preceded and activated by Ser/Thr kinase (Raf) rather than by a tyrosine kinase, a burst of interest in dual kinases led to a rapid increase in the number of enzymes reported to display this feature (e.g., see Refs. [36–41]). Though the number of 'dual kinases' still keeps rising, the significance of their apparent dual specificity has been drastically reassessed. In some instances the dual specificity is detectable only as autophosphorylation while the canonical kinase activity, tested with *trans* substrates, is uniquely directed toward either tyrosyl (e.g. [42]) or seryl (e.g. [38]) residues. In other cases the practical relevance of tyrosine phosphorylation is dubious since the K_{cat} values with tyrosyl substrates are negligible compared to those with canonical Ser/Thr substrates (e.g. [37,40]). It should be recalled in this connection that even bona fide Ser/Thr kinases like phosphorylase kinase [43], Spk1, calcium/calmodulin-dependent protein kinase II and PKA [43a] can exhibit tyrosine kinase activity under special conditions and a canonical tyrosine kinase like Src can phosphorylate aliphatic as well as aromatic alcohols in a stretch of peptide [22]. It is likely therefore that many of these are 'ambiguous' rather than 'dual' in behaviour, due to artefacts. They may have some mechanistic interest, but their functional relevance is dubious. The only physiologically relevant example of dual specificity protein kinase re-

mains that of the MAP kinase kinase (MEK) family [44–46]. These are not pleiotropic kinases, however, impinging on a variety of substrates where they might phosphorylate either Ser/Thr or Tyr sites, but highly dedicated kinases committed to the simultaneous phosphorylation of *both* Thr and Tyr in a conserved sequence, T-X-Y, of the activation loop of MAP kinase and totally inactive on a tryptic peptide [47] or on a 22 residue synthetic peptide (Vaglio et al., unpublished) including the same site.

2.4. Positive and negative specificity determinants

Extensive studies on a wide variety of protein kinases have clearly demonstrated that the primary sequence immediately surrounding the phosphorylated residues of the substrate plays an essential role in substrate recognition. The elements of this sequence that are especially required to ensure efficient phosphorylation are referred to as *specificity determinants*. Both the nature and the position of specificity determinants depend on the kinase considered; sometimes the replacement of a crucial determinant at the canonical position is tolerated if it is present at subsidiary positions. The relevance of individual determinants can also vary, suggesting that some of them are 'indispensable', while others are 'subsidiary', although it is not always easy to draw a borderline between these two categories. Negative determinants, whose presence can compromise the phosphorylation of otherwise suitable sites, should also be taken into account. While this deleterious effect is well documented in some cases, little is known about the negative determinants of most protein kinases.

3. Specificity determinants of Ser / Thr protein kinases

Both from the inspection of sequences around phosphorylated residues and from studies with synthetic peptide substrates it has become clear that Ser/Thr protein kinases are, with few exceptions, markedly 'sequence specific'. The residues utilized for site recognition by these kinases fall into five categories: (a) basic residues (normally Arg and/or Lys); (b) prolyl residues; (c) carboxylic residues (Glu and Asp); (d) hydrophobic residues; (e) previously phosphorylated seryl, threonyl and tyrosyl residues. According to their different preferences for these types of determinants a rough classification of Ser/Thr kinases into 3 main groups can be made: (i) basophilic kinases, using basic and often also hydrophobic residues, as determinants; (ii) proline-directed kinases, some of which also require basic residues, besides proline and (iii) acidophilic/phosphate-directed kinases, with carboxylic and phosphorylated residues variably interchangeable.

3.1. Basophilic Ser / Thr protein kinases

Before the wide distribution of some acidophilic and Pro-directed protein kinases was recognized the requirement of basic specificity determinants was held as a general feature of all Ser/Thr protein kinases. The majority of known Ser/Thr protein kinases still fall into this category, including the whole of the large ACG group (PKA, PKG, PKC, RSK kinases etc) with a couple of exceptions, all the members of the CaMK group (also inclusive of AMP kinases) and a number of representatives of kinases not falling into major groups according to the Hanks and Hunter nomenclature [25]. Often the basic determinants are situated on the N-terminal side of the target residue, but sometimes they are found on both sides and seldom only C-terminal. In general arginine is preferred over lysine, but notable exceptions are also known. Quite frequently hydrophobic determinants are also recognized by basophilic kinases, and sometimes the role of hydrophobic side chains is actually predominant over that of basic residues. Conversely a prolyl residue at position $n + 1$, the hallmark of Pro-directed kinases, has proved detrimental whenever tested with kinases of this group.

3.2. Pro-directed Ser / Thr protein kinases

This growing group of Ser/Thr kinases is composed of those members of the CMGC group that belong to the families of cyclin-dependent kinases and of MAP kinases (ERKs), also including JNK, SAPKs and RK. All these kinases share the stringent requirement for a Pro at position $n + 1$, a feature that, in contrast, is detrimental for most protein kinases and for many protein phosphatases as well (see

Section 7). They are, however, distinguished by the additional preference of MAP kinases, but not of cyclin-dependent kinase, for a second proline, and by the requirement of cyclin-dependent kinases, but not MAP kinases, for basic residues.

3.3. Acidophilic and phosphate-directed Ser/Thr protein kinases

Only few Ser/Thr protein kinases utilize carboxylic and/or phosphorylated side chains as specificity determinants. Some of them however are ubiquitous and pleiotropic enzymes whose substrates are numerous and widespread. Acidophilic protein kinases belong to quite distinct groups of protein kinases, according to the Hanks and Hunter nomenclature [25]: CK2 ('casein kinase-2 or -II') and GSK3 constitute two distinct branches of the CMGC group, also including the cyclin-dependent and the mitogen-activated (MAP) kinases, both Pro-directed. βARK and rhodopsin kinase are members of the ACG group, mainly composed of basophilic kinases, and CK1 ('casein kinase-1 or I') is a growing family of enzymes not falling into any one of the major groups according to the Hanks and Hunter classification. Finally it should be mentioned that acidophilic Ser/Thr protein kinases also include the genuine casein kinase(s) (whose primary structure is still unknown) that phosphorylate casein in the lactating mammary gland and a putative family of secretory pathway protein kinases sharing with the mammary gland casein kinase(s) the ability to phosphorylate Ser-X-Glu motifs. GSK3 and CK1 are considered primarily phosphate-directed enzymes and with GSK3 there is no evidence that carboxylic residues can substitute for phosphoserine; in contrast, CK2 and the 'casein kinases' are believed in most instances to be Glu/Asp and Glu directed kinases, respectively, although phosphorylated residues can sometimes replace the carboxylic ones with variable efficiency.

4. Specificity determinants of protein tyrosine kinases (PTKs)

The definition of the specificity determinants of PTKs has not been as clear-cut as that of Ser/Thr protein kinases. Early work indicated that acidic residues on the N-terminal side of tyrosine were positive determinants for these kinases. However attempts to obtain efficient peptide substrates based on these residues were not very successful and often the substitution of these residues by neutral ones was tolerated fairly well by a number of tyrosine kinases, especially those of the Src family. Consistent with this is the observation that several tyrosyl phosphoacceptor sites lack acidic residues in the proximity of the N-terminal side of tyrosine (e.g., Refs. [48–50]). Where present, moreover, the acidic residues are not situated at consistently fixed positions. These observations suggest that other factors are more important than the sequence of the phosphoacceptor site for PTK specificity. Several arguments make this expectation plausible, among which are the ability of receptor PTKs to recruit their substrates by autophosphorylation followed by interaction with the SH2 domains of many substrates and the presence in non-receptor PTKs of modules, like the SH2 and SH3 domains, that mediate interactions with potential targets independently of their catalytic domains. Even the N-terminal 'unique' domain of non-receptor PTKs has been shown to cooperate in substrate selectivity [51], while a new module, termed PTB or PID, capable of recognizing phosphotyrosyl residues but distinct from the SH2 domains has recently been discovered [52–54]. A more indirect recruitment of PTK targets can be mediated by the pleckstrin homology (PH) domains [4,5] which interact with membrane phospholipids and may thus facilitate the juxtaposition of phosphorylatable proteins with the membrane-bound PTKs [55]. In conclusion, it seems likely that substrate recruitment by PTKs is mediated by multiple factors, a circumstance that makes direct site recognition less critical than in the case of Ser/Thr protein kinases. This difference may also be reflected in the small number of specificity determinants recognized by PTKs, which seem to be restricted to just acidic and hydrophobic (and/or β-branched chain) residues (see Section 5.3).

It should be stressed, however, that the fidelity of signal transduction via protein tyrosine kinases also depends critically on the sequence specificity of their catalytic domain supporting the view that the consensus sequences play a crucial role in the case of tyrosine phosphorylation as well [12]. Consistent with this is the ability of EGF receptor protein kinase to

phosphorylate at specific residues two proteins lacking SH2 domains and therefore unable to dock to the receptor autophosphorylation site [56]. These observations highlight the importance of unraveling the differences in sequence specificity of PTKs as well, although it is clear that additional features possibly residing outside the phosphoacceptor site may deeply affect the targeting by this class of protein kinases.

5. The consensus sequence: an elusive concept

Theoretically the consensus sequence of a given kinase is obtained by placing the specificity determinants outlined above at the positions, relative to the target residue, where they are maximally effective. This would also provide a general picture of the primary structure of any site phosphorylatable with 'reasonable' efficiency by that kinase. According to the convention generally adopted [2] the consensus sequence should include the target aminoacid (position $n \pm 0$) and specify the positions relative to it ($n+1$, $+2$, $+3.../n-1$, -2, etc.) of the crucial specificity determinants (indicated by their one letter code), whereas an X denotes those positions where any residue would be in principle tolerated, with the implicit exclusion, however, of those which act as negative determinants (most of which are still unknown). Where two or more aminoacids function interchangeably these are listed at a given position with a slash separating them.

Clearly, to establish a simple and reliable consensus sequence is always a laborious task whose outcome may be more or less satisfactory depending on the amount and quality of data available and on the interpretation of terms like 'reasonable' phosphorylation and 'interchangeable' residues (which incidentally can be 'interchangeable' at certain positions but not at others). It has first to be emphasized in this respect that while the specificity of some protein kinases (e.g., PKA, CaM kinase II, PKC, CK2, AMPK, MAPKAP kinases) has been thoroughly studied with the aid of many substituted peptide substrates and by different approaches, little or no information is available about many other kinases. The shortage of information is even more evident for *negative* determinants. Second, the borderline between phosphorylatable and non-phosphorylatable substrates, necessary for defining a 'minimum requirement' consensus sequence is rather elusory: by increasing the concentration of the substrates and/or of the kinases one can usually detect the phosphorylation of nearly any peptide by nearly any kinase. At the other extreme, the definition of an 'optimal' sequence, ensuring maximal phosphorylation efficiency, is nearly impossible, as it is hard to dismiss the possibility that larger and/or peptides of different sequences may exist which are phosphorylated more efficiently that the best available one. It may also be pertinent to ask whether such an optimal peptide substrate would really reflect a physiological situation, considering that in some instances the artificial mutation of natural phosphoacceptor sites improves their phosphorylation over that of the wild-type substrate. On the other hand, crucial determinants established with model peptide substrates may prove less important in the intact protein substrate, as exemplified by MLCK [57,58].

In general, consensus sequences represent a compromise between sequences ensuring 'minimum' and 'optimal' phosphorylation efficiencies. In particular those determinants are taken into account whose individual replacement causes dramatic drops in efficiency, either by increasing the K_m or decreasing the Kcat by one or more orders of magnitude. Even this criterion, however, has no absolute validity, since the importance of an individual residue at a given position may vary greatly depending on the structure

Table 1
Protein kinase specificity at certain positions is influenced by the rest of the sequence

Kinase	Peptides	V_{max} (relative)	K_m (μM)	Ref
CK2	**S**EEEEE	100	270	[98]
	SE**A**EEE	74	1950	
	RRRADD**S**DDDDD	100	15	[88]
	RRRADD**S**D**A**DDD	66	9	
cdc2	NFK**S**PVK**T**IR	100	43	[106]
	NFK**S**PV**A**TIR	< 10	n.d.	
	SPPRRRRK	100	5.0	[58a]
	SPP**A**RRRK	67	6.3	

The phosphorylated residues are underlined; residues which replace putative specificity deteminants are bold type and underlined. n.d., not determined due to too low phosphorylation rate.

surrounding it, as shown in Table 1 for an acidic residue at $n+2$ and a basic residue at $n+3$ with CK2 and cdc2, respectively. These and other similar observations show that sometimes important determinants may become dispensable if 'subsidiary' positive determinants are located elsewhere. These 'substitutive determinants' in the case of CK2 and MAP-KAP kinase 1 have been shown to generate atypical consensus sequences lacking the 'canonical' determinants at position $n+3$ [59] and $n-5$ [31], respectively. They may also account for some discrepancies between conventional consensus sequences and those constructed with the data provided by degenerate peptide libraries whose employment for the definition of protein kinase specificity will be discussed in the following chapter.

5.1. Searching for 'optimal sequences' with peptide libraries

All the consensus sequences considered in previous reviews on protein kinase specificity (e.g., Refs. [1,2]) were obtained by a uniform protocol consisting in the inspection of the sites phosphorylated by a given kinase followed by assays with individual peptide substrates reproducing the natural sites with suitable modifications of residues that were suspected to play a role in substrate recognition.

Recently, a new approach has been developed which is based on the synthesis, either by chemical [9–11] or cassette mutagenesis [8] methods, of fully or partially degenerate peptide libraries composed of myriad constituents varying in aminoacyl residues at every or definite positions. These libraries are subjected as such to phosphorylation by a selected kinase, the phosphorylated peptides are separated from the non-phosphorylated ones by various devices and finally they are sequenced either together (i.e., as a mixture of phosphorylated peptides) [11] or as single entities [10]. The former approach is only possible if the peptide library is 'oriented', i.e. it contains the selected phosphorylatable residue at a fixed position and it is not fully degenerate at other positions, where phosphorylatable residues are absent. While, as discussed above, the conventional approach is mainly aimed at outlining the minimum structural requirements of kinases, the data from the libraries are expected to provide 'optimal' sequences within the limits of the preselected peptide length and degeneracy. The method that has been more widely applied up to now is based on oriented libraries consisting of peptides including either a seryl [11] or a tyrosyl [12] residue embedded in the middle if a sequence of 8 degenerate residues, thus encompassing positions from $n-4$ to $n+4$. After phosphorylation by the kinase and isolation of the phosphorylated species, these are sequenced together and the relative amount of each residue after every cycle will reflect the selectivity for that residue at a given position. In principle the sequence constructed from the residues displaying the highest selection at each position would represent the optimal sequence recognized by the kinase, but a complication may arise from the observation (Section 5) that the relevance of a residue at a given position also depends on the rest of the sequence (see Table 1). Data obtained with this library using PKA, $p34^{cdc2}$ and $p33^{cdk2}$ [11] are essentially in agreement with the information provided by the inspection of the phosphoacceptor sites and by conventional peptide substrates. A fairly good correlation also exists between the optimal sequences found for nine protein tyrosine kinases and the local features of many tyrosine phosphoacceptor sites [12]. Some discrepancies between these data and conclusions drawn from more conventional approaches and another random library approach [13] are discussed below (Section 5.3.2).

An outstanding advantage of the peptide libraries is that they can provide with a single experiment an amount of information that would otherwise require hundreds of experiments with a huge variety of successively designed peptide substrates. They will also prove very helpful for the quick scrutiny of the specificity of protein kinases whose substrates are unknown and perhaps for designing artificial peptides suitable for monitoring the activity of protein kinases which are inactive on peptides reproducing their natural targets.

A limitation of the oriented libraries, although one that can be obviated, is the relatively modest extension of the two 'wings', excluding possible determinants eventually acting at positions further away. A structural constraint is also imposed to the oriented libraries by the fixed position of the phosphorylatable residue, which in some instances could hamper the detection of situations more favourable for phospho-

rylation, where the phosphate acceptor residue might be located at different positions (see Section 5.3.2).

In principle the data provided by peptide libraries could be compromised by the amazingly low concentration of each peptide in the assay medium. It can be calculated that in a XXXX-S-XXXX oriented library, with a theoretical degeneracy of 15^8 and tested at a concentration of 3.3 mg/ml [11], the concentration of each peptide species would be about 1.2 pM, i.e. several orders of magnitude below the K_m of the best known peptide substrates for protein kinases (around 1 μM). Such a figure dramatically favours peptides displaying extremely low K_m values, while abolishing differences that may arise from variable K_{cat} values alone. Generally speaking this is more likely to be an advantage rather than a drawback, since affinity is the most important feature in determining protein kinase selectivity. However an exaggerated emphasis on affinity compared to catalytic rate could lead to misleading conclusions whenever peptides exhibiting low but variable K_m values are scrutinized. Clearly peptides exhibiting the very lowest K_m values will be selected irrespective of the fact that other peptides might be 'better' substrates for displaying higher phosphorylation rates at 'physiological' concentrations.

The optimal sequence provided by the oriented library approach established by Cantley and collegues [11] is merely a virtual entity, resulting from the simultaneous analysis of a variety of phosphorylated peptides, each one contributing to the 'overall' sequence in proportion to its relative abundance in the mixture. This may lead to overestimating the importance of positive determinants at positions other than

Table 2
Optimal sequences provided by oriented peptide libraries may be inconsistent with consensus sequence

	V_{max}	K_m (μM)	V_{max}/K_m	Ref.
A.				
RRR**R**SVA	22.5	8	2.75	[65]
RRR**A**SVA	19.5	6	3.25	[65]
Consensus:				
RR**X**S				
B.				
Library:	XXXXSXXXX	selection at position $n-1$		
	R	6.5		[11]
	A	0.5		[11]
optimal sequence (virtual):				
RR**R**S				[11]
C.				
	phosphopeptides from the library sequenced together (X = any residue other than R)			
optimal sequences:	-R-R-(X/R)-S-			
other sequences:	-R-X-R-S-			
	-X-X-R-S-			
	-X-R-R-S-			
Predominant residue after cycle	1 2 3			
	R			
	R			
	R			
resulting sequence:	RRRS			

A. The phosphorylation efficiency of a peptide fulfilling the consensus sequence for PKA is not reduced by replacing Arg at $n-1$ by Ala. B. Arg at position $n-1$ displays a much higher selection than Ala according to the oriented library approach. C. The apparent discrepancy between A) and B) can be explained considering that the yield of Arg at $n-1$ is higher than that of Ala (though they are equivalent in the consensus sequence peptides) if the whole set of phosphopeptides are simultaneously sequenced. For additional explanation see text.

Table 3
Sequence specificity of protein kinases

	Consensus sequence/*phosphorylated sequence* [b]	Ref.
Ser/Thr kinases [a]		
PKA	**R**-(**R**/K)-X-**S**/**T**-**B**	see text
PKG	**R**-(R/K)-X-**S**/**T**-**B**	see text
PKC	(R/K)-(**R**/**K**)-X-**S**/**T**-**B**-(**R**/**K**)-(**R**/**K**)	see text
	$(R/K)_{1-3}$-$(X)_{2-0}$-**S**/**T**-$(X)_{2-0}$-$(R/K)_{1-3}$	[2]
	X-**R**-X-X-**S**/**T**-X-**R**-X	[75]
PRK1	*KKKKKRFSFKKSFKLSGFSFKKSKK*	[79]
PKB/RAC	**R**-X-**S**/**T**-X-**K**	[80]
	GRARTSSFAE	[81]
	GRPRTTSFAE	[81]
$p70^{s6k}$	(**K**/**R**)-X-**R**-X-X-**S**/**T**-B	[31,82]
$p90^{s6k}$/MAPKAP-kin.1	X-X-(**R**/**K**)-X-**R**-X-X-**S**-X-X	[31]
	R-**R**-**R**-X-**S**-X-X	[31]
MAPKAP kin.2	X-X-**B**-X-**R**-X-X-**S**-X-X	[84]
CaM kinase I	**B**-X-**R**-X-X-**S**/**T**-X-X-X-**B**	[32,90]
CaM kinase II	**B**-X-(**R**/**K**)-X-X-**S**/**T**-X-X	[84]
CaM Kinase III (= EF-2 kinase)	*AGETRFTDTRK*	[151]
Cam kinase IV (= Cam kin. Gr)	*VPGKARKKSSCQLL*	[152]
	KSDGGVKKRKSSSS	[153]
	PLARTLSVAGLP	[153]
MLCK (smooth m.)	K-K-**R**-X-X-**S**-X-B-B	see text
PhK	**K**-R-**K**-Q-I-**S**-**V**-**R**	see text
AMP-K	**B**-(X, **R**/K/H)-X-X-**S**/**T**-X-X-X-**B**	[32,154]
HMG-CoA Red. kin. A	**B**-(X, **R**/K/H)-X-X-**S**-X-X-X-**B**	[32]
SNF-1 PK (yeast)	**B**-X-**R**-X-X-**S**-X-X-X-**B**	[32]
Autophosph. dep. PK	R-X-(X)-**S**/**T**-X-X-X-S/T	[155]
PAK-1 (cardiolipin activated PK)	*PLSRTLSVAA*	[156]
	AKRRRLSSLRA	
	VRKRTLRRL	
HRI(HCR)	(E)-X-**S**-X-**R**-X-X-**R**	see text
PKR(dsRNA kin / DAI)	(Rn)-X-X-**S**/**T**-X-**R**-X-X-**R**	see text
NIMA	F-R-X-**S**/**T**	[157]
	[R-**F**-R/K-R/K-**S**-R/I-R/I-M-I]	c
Nek 2	*IRRLSTRRR*	[158]
HSV and PRV PKs	**R**-**R**-**R**-**R**-X-**S**/**T**-X	see text and [65]
TGFβ rec. PK	K-K-K-K-K-K-**S**/**T**-X-X-X	[14]
Bck1 (SLK, Ssp31)	[R-R-F-G-**S**-B-R-R-F]	[11]
ERK (MAP kin.)	P-X-**S**/**T**-P-P	[159]
	[X-X-**P**-X-**S**/**T**-**P**-**P**-P-X]	d
cdc2 PK	X-**S**/**T**-**P**-X-**K**/**R**	see text
	[H-H-H-K-**S**-**P**-R-**R**/**K**-R]	[11]
cdk2	[H-H-H-R-**S**-**P**-R-**K**-R]	[11]
cdc2 like PK	P-X-**S**/**T**-**P**-K-**K**-X-K-K	[160]
Ceramide act. PK	PTLP > PLTLP > PLLTP = PLTP	[94]
PITSLRE PK	*SPMRSRSPSRSK*	[161]

Table 3 (continued)

	Consensus sequence/*phosphorylated sequence* [b]	Ref.
CK2	**X-S/T-X-X-(E/D/Sp/Yp)-X**	
	X preferably acidic	see text and [103]
	(Sp/D)-(Sp/D)-S-(Sp/D)	see text and [59]
Golgi casein kinase(s)		
(GEF-CK)	**X-S-X-(E/Sp)**	see text
CK1	**(Sp/Tp)**-X-X-(X)-**S/T-B**	see text
	$(D/E)_{\geq 4}$-X-X-**S/T**-B	see text
GSK3	**S/T**-X-X-X-**Sp**-X	see text
	(X,P)-**S/T**-(X,P)?	see text
βARK	$(D/E)_n$-**S/T**-X-X-X	[128]
RhK	X-X-**S/T**-$(E)_n$	[128]
	X-**S/T**-X-X-X-(A/P/S/T)	[162]
T-389 $p70^{s6k}$		
putative kinase	VFLGF*T*YVAP	[130]
PDHK	Xn-**S**-X-X-**D**-X-X	[131]
BCKDHK	*HSTSDD and YRSVDE*	[132]
DNA-dependent PK	X-S-Q	[135]
	P-S/T-X	[136]
Tyr-kinases		
Syk (p72/p38)	$(E/D/Tp/Yp)_{1-3}$-X_{0-2}-**Y**-E-E-B	see text
c-Src	[Y-**I-Y**-G-S-F-K] [e]	[13]
v-Src	*KKKEEEEEEEYMPMEDL*	[163]
	[E-E-E-**I-Y**-(G/E)-E-**F**-D]	[48]
Lck	[X-E-X-**I-Y**-G-V-L-F]	[48]
c-Fgr	[M-E-(E/N)-**I/V-Y**-(G/E)-I-F-F]	d
Lyn	[**D**-E-E-**I-Y**-(E/G)-E-L-X]	d
Abl	*KKKEEEEEEEYMPMEDL*	[163]
	RRLIEDAIYAARG	[9]
	[?-X-V-**I-Y**A-A-P-F]	[48]
c-Fps/Fes	[E-E-E-**I-Y-E**-E-I-E]	[48]
Fak	*EEEHVYSF*	[164]
CSK	[X-X-X-**I-Y**-(M/I/F/)-F-F-F]	d
EGF Receptor	[E-E-E-E-**Y**-F-E-**L**-V]	[48]
PDGF Receptor	[E-E-E-E-**Y**-V-F-I-X]	[48]
FGF Receptor	[?-E-E-E-**Y**F-F-L-F]	[48]
Insulin Receptor	[X-E-E-E-**Y**-M-M-M-M]	[48]
	KKSRGDYMTMQIG	[149,150]
	KKKLPATGDYMNMSPVGD	[149,150]
IGF1 receptor	*KKKSPGEYVNIEFG*	[150]

the canonical ones, through the phenomenon of 'substitutive determinants' discussed above (Section 5). An illuminating example is provided by PKA. This kinase normally recognizes two crucial arginyl residues at positions $n-2$ and $n-3$. Whenever this condition is not fulfilled the resulting low phosphorylation efficiency can be improved by arginine(s) at positions other than -2 and -3, notably at -4 to -7 (Section 5.2.1) and -1. If, however, the crucial determinants are present, then an additional arginine at $n-1$ is totally unimportant, an alanine being equally effective (see Table 2A). Nevertheless the information provided by the oriented library is that the optimal sequence is RR*R*S, with arginine at $n-1$

position being selected much more strongly that alanine (Table 2B). It is likely that this discrepancy arises from the fact that during the sequencing cycle that removes the $n-1$ residues the *optimal* peptides present in the mixture will give rise to approximately the same amount of R and A, but *all the other* peptides are richer in arginine than in alanine to compensate the lack of arginine at the canonical positions. Thus the *calculated* optimal sequence (Table 2C) will emerge as RR**R**S, although the *real* sequence of the *optimal* peptides is just RR**X**S (X not being preferentially a basic residue). The correctness of this interpretation is corroborated using 'spatially addressable peptide libraries' that allow the identification of *individual* optimal peptides, instead of analysing a mixture of phosphopeptides. In this case arginine is less selected than alanine at position n-1 [14,15]. It can be predicted that this feature of the oriented library approach will give rise to 'expanded consensus sequences' every time a kinase recognizes specificity determinants at subsidiary positions if they are lacking at the canonical one(s) as has been shown, e.g., of CK2 and cdc2 (see Table 1) and of some protein tyrosine kinases [60]. In these cases the predilection for specificity determinants at positions other than the canonical ones will be overestimated by oriented peptide libraries.

5.2. Consensus sequences of individual Ser / Thr protein kinases

A list of consensus sequences recognized by individual protein kinases (or groups of protein kinases) is shown in Table 3. As mentioned before, they are distilled from data in the literature that may vary strikingly from one kinase to another both in terms of thoroughness and criteria adopted. In a few cases, in the absence of any systematic analysis of putative determinants, just the sequence(s) phosphorylated by a given kinase are presented (in *italics*) with those residues that have been suggested to play a role in recognition in bold type.

Wherever available, the information provided by protein substrate mutagenesis and by the novel peptide library methods has been taken into account. Attention has also been given to the structural features of protein kinases that could account for site selectivity (see Section 9.2). Wherever applicable we have tried to discriminate between essential determinants (bold type) and accessory residues that simply improve phosphorylation efficiency. Slash(es) separate interchangeable residues without denoting any order of preference; thus, residues separated by slashes may either be similar or different in their effectiveness. Residues that are likely to be interchangeable (e.g., D instead of E or K instead of R) but whose effect has not been checked (at least to our knowledge) are normally *not* indicated. 'B' however denotes any hydrophobic residue, although only some of them may have actually been found and/or tested at the indicated positions. 'X' denotes any residue, with the important assumption, however, that it is not a negative determinant at that given position.

The consensus sequences and, more generally, the specificity determinants of some important types of protein kinases are reviewed briefly in the following commentary.

Notes to Table 3:

[a] The acronyms used to indicate protein kinases are the same adopted by Hanks and Hunter [25] (see 'Abbreviations'). Protein kinases indicated in *italics* are dedicated enzymes whose phosphorylation efficiency with peptide substrates is invariably very low.

[b] Aminoacids are indicated by the one letter code; B stands for any hydrophobic aminoacid and X for any residue except those which may play a negative role (see text). Sp, Tp and Yp denote phosphoserine, phosphothreonine and phosphotyrosine, respectively. Typical consensus sequences are written with hyphens separating aminoacids; interchangeable residues at a given position are grouped between parentheses, (), and separated by slashes (see also text). Sequences entirely bracketed, [ ], refer to optimal virtual sequences resulting from sequencing the peptides phosphorylated in oriented peptide libraries (see text). Bold type denotes residues that are known or suspected to play a crucial role as specificity determinants and/or are strongly selected in peptide libraries. The target residues are both bold type and underlined. Sequences indicated in *italics*, without hyphens, denote peptides or protein sites that have been shown to be efficiently phosphorylated by kinases whose site specifity could not be unambiguously clarified yet.

[c] Songyang and Cantley, personal communication.

[d] Ruzzene et al., unpublished.

[e] Optimal peptide substrate for c-Src identified within a fully randomized library of heptapeptides [13].

5.2.1. PKA

cAMP-dependent protein kinase (PKA) was the first protein kinase to undergo specificity studies based on both the inspection of phosphoacceptor sites and the design of model peptide substrates [61–64]. The crucial importance of basic residues on the N-terminal side of serine was immediately perceived and led to the definition of PKA consensus as being XRRXSX [1]. These studies also suggested the possible importance of hydrophobic residues at $n+1$ position. This concept was corroborated by more recent observations with either substituted peptides [65, 66] or an oriented peptide library [11] and support the validity of a phosphorylation motif including this determinant as well (XRRXSB), which is also consistent with the crystal structure of the binary complex of PKA catalytic subunit and an inhibitor PKI peptide [67]. The crucial relevance of this hydrophobic determinant was also implicit in the finding that the serine exhibiting the motif SI is preferred over the one fulfilling the canonical motif RRXS in the peptide VLQRRRPS*S*IPQ [68]. Lysine cannot replace arginine as an efficient specificity determinant [11,65]; threonine is a much poorer phosphate acceptor than serine, and proline is a strong negative determinant if present at position $n+1$ [69], a finding also in agreement with the negative selection of proline at this position in a peptide library [11]. Nevertheless a few physiological phosphoacceptor sites for PKA (notably in protein phosphatase inhibitor-1 and in its homologue DARPP-32) exhibit the motif -TP-. It is also clear that arginine is helpful at positions other than the canonical ones, especially if it is absent at these latter positions. In particular the peptide RRRR*AA*ASVA is almost as good as the canonical substrate *RR*ASVA [65]. Compensatory effects like this could account for the persistent phosphorylation of PKA substrates where all the canonical sites have been ablated [70] and for the very efficient phosphorylation of the β-subunit of phosphorylase kinase at a site where arginine residues are located at $n-2$ and $n-6$ and not at $n-3$.

5.2.2. PKG

The sequence specificity of the cGMP dependent protein kinase is very similar to that of PKA, consistent with the high similarity of the catalytic domain, including the conservation of the residues interacting with the basic residues at positions $n-2$ and $n-3$ and the hydrophobic character of binding determinants for residue at $n+1$. However it is not identical [15,71]; in particular the arginine at position $n-2$ is less important than that at $n-3$ [66] and it appears that PKG requires more basic residues than PKA [15]. In sharp contrast to the similarity in sequence specificity, PKG displays quite a different 'active site specificity' as disclosed using alcohol-containing peptide substrates [24]. This difference has been exploited for synthesizing a specific inhibitor of PKG which is ineffective on PKA [72]; it could also account for the finding that the G-substrate, one of the best substrates of PKG, is phosphorylated at threonine. In peptide substrates, however, PKG, similar to PKA much prefers serine over threonine as phosphoacceptor residue [71].

5.2.3. PKC

PKC was originally thought to share the site specificity of PKA, and the 'kemptide' (originally designed for PKA) was routinely used for its assay. Later it was show that the basic determinants can be located on the C-terminal side of serine as well [73]. PKC proved especially active on seryl residues embedded between two clusters of basic residues, which in some instances are unaffected by PKA [74]. Such an ambivalent basophilic nature of PKC is also in agreement with predictions based on the structure of its catalytic domain in which the acidic residues responsible for upstream basic recognition by PKA are conserved, whereas basic residues at the end of subdomain II which are responsible for the acidophilic nature of CK2 are replaced by acidic ones in PKC (see Section 9). The hydrophobic character of PKA binding determinants for residue at position $n+1$ is conserved in PKC, consistent with the favourable role of a hydrophobic residue at this position in PKC as well [74a]. It is possible, however, that the relevance of this feature may vary depending on the presence of basic residues on both or just on one side of the phosphorylatable aminoacid. While in fact optimal phosphorylation requires basic residues on both sides [74a,75], their presence only either upstream or downstream is also tolerated, as reviewed in [2]. In this connection it should be remembered that the numerous isoforms of PKC (no less than 11 are known, grouped into 'conventional',

'novel' and 'atypical' PKCs) may display significantly different specificity requirements (referenced in [76]). In particular PKCδ seems to be the most tolerant one as far as the number of basic residues and their absence downstream are concerned: it is the only isoform that appreciably phosphorylates Thr-431 in eEF1α and in the derived peptide RFAVRDMRQ*T*VAVGVIKAVDKK [76]. In contrast, derivatives including additional basic residues at position $n-3$ and $n+2/n+3$ are readily phosphorylated by the other isoforms as well. A remarkable feature of PKC is its ability to recognize lysine as efficiently as arginine [65]. Moreover its preference for Ser over Thr is not so marked as in the case of PKA [73,77].

The specificity of recently discovered *P*KC *r*elated protein *k*inases (PRKs) [78] has not been investigated. Precise overlapping with PKC in the phosphorylation of multiple residues of MARCKS [79] would suggest very similar specificity to PKC.

5.2.4. PKB (RAC)

Both acronyms refer to the relatedness of this kinase (the cellular homolog of the oncogenic form v-Akt) to both PK*A* and PK*C*. Its main phosphorylation site in MBP (*R*GS*S*G*K*D) and data with model peptides indicating that basic residues are required on both sides of serine [80] suggest closer similarity to PKC specificity. However, the sites affected by RAC in two isoforms of GSK3 (GRARTS*S*FAE and GRPRTT*S*FAE) [81] do not conform to this prediction. Clearly, additional investigation will be required in order to define the consensus sequence of PKB.

5.2.5. Ribosomal protein S6 kinases and MAPKAP kinases

Only the smaller members ($p70^{s6k}$) of this group appear to be bona fide S6 protein kinases, while the larger ones ($p90^{s6k}$/MAPKAP kinase 1) are probably not implicated in S6 phosphorylation in vivo although they do phosphorylate it in vitro. Both utilize basic residues on the N-terminal side of the target residue as major specificity determinants [82] and the motif xxR/KxRxxSxx is recognized by both kinases equally well [31]. Unlike $p70^{s6k}$ however, $p90^{s6k}$/MAPKAP kinase-1 tolerates the lack of R/K at $n-5$ if arginine is present at $n-2$ and $n-3$, in the alternative motif, xxRRxS. Moreover $p70^{s6k}$ differs from $p90^{s6k}$ in accepting Thr instead of Ser and in tolerating the absence of any residue at $n+2$ in peptide substrates [31]. These properties have been exploited in designing a specific peptide substrate which is almost unaffected by $p90^{s6k}$ [31]. Although at the crucial $n-5$ position Arg and Lys are interchangeable, it appears that at other positions Lys is minimally effective compared to Arg, especially with $p90^{s6k}$/MAPKAPK-1, formerly termed ISPK-1 [83]. This kinase also appears to tolerate a non-hydrophobic residue at $n+1$ [83], while proline at the same position is detrimental with both $p70^{s6k}$ and $p90^{s6k}$/MAPKAPK-1 [31,83]. The ordered phosphorylation mechanism of S6 protein by $p70^{s6k}$ is consistent with the consensus sequence shown in Table 3. The first and second residues undergoing phosphorylation (Ser236 and Ser235, respectively) conform to the (**K/R**)-X-**R**-X-X-*S*-(B) pattern, though the second one lacks the dispensable hydrophobic residue at $n+1$. The subsequent phosphorylation of the downstream Ser240, Ser244 and Ser247, on the other hand, may proceed independently of stringent recognition elements since the kinase-substrate complex does not dissociate till the progressive phosphorylation is completed [82].

MAPKAP kinase-2, whose catalytic domain is more closely related to calmodulin-dependent protein kinases than to S6 protein kinases, neatly differs in site specificity from both $p70^{s6k}$ and $p90^{s6k}$, although all three members of the family utilize a common recognition pattern, iuiuuS, where i and u stand for 'important' and 'unimportant' residue, respectively. The very important residue recognized by MAPKAP kinase-2 at position $n-5$ is hydrophobic, however, rather than basic [84]. In both MAPKAP kinases the replacement of a serine by a threonine is extremely unfavourable.

5.2.6. Calmodulin-dependent protein kinases

The common denominator of the members of this otherwise rather heterogeneous family of kinases is their susceptibility to activation by calmodulin. The inspection of their phosphorylation sites would suggest that they are all more or less basophilic. However, on the basis of substrate specificity they can be clearly divided into two categories: those with broad specificity (CaM kinase II, CaM kinase IV and probably CaM kinase I) which serve a multiplicity of physiological roles [85] recognizing canonical con-

sensus sequences in a variety of protein targets, and those with restricted specificity (phosphorylase kinase, myosin light chain kinase and CaM kinase III), whose targeting mainly depends on structural features outside the phosphoacceptor site [86]. The definition of the consensus sequence of these latter dedicated CaM kinases, which normally phosphorylate peptide substrates with very low efficiency, is not only problematic but even of little use since it is unlikely that these enzymes will recognize these motifs in proteins other than their targets. In the case of phosphorylase kinase, however, recent studies with mutated enzyme [87] and comparative analysis of its sequence with those of kinases whose substrate recognition elements have been identified [67,88] support the outcome of pioneering studies [89] that two basic residues at positions $n-3$ and $n+2$ and a hydrophobic residue at $n+1$ play a definite role in substrate recognition; in contrast other basic residues present in the physiological substrate phosphorylase-b upstream from $n-3$ seem to play a minor role. Using a similar mutational approach the importance of the three basic residues at positions $n-6$ to $n-8$ for recognition by MLCK was redefined [57,58,86] while that of arginine at $n-3$ and of hydrophobic residues at $n+1$ to $n+3$ was corroborated.

The pleiotropic CaM kinases I and II (this latter also termed multifunctional) conform to the same recognition pattern, iuiuuS of S6 protein kinases (see Section 5.2.5), they differ however in accepting threonine and serine equally. They also show differences in recognition pattern, in that, e.g. lysine instead of arginine at position $n-3$ is better tolerated by CaM kinase II [84] than by CaM kinase I [32,90].

5.2.7. AMP-activated / SNF1 protein kinases

The members of this family of protein kinases activated in response to stress and starvation are structurally, albeit not functionally, related to the calmodulin-dependent protein kinases and they recognize similar motifs on synthetic peptide substrates of the type BxRxxS/TxxxB (where B stands for hydrophobic), recalling the recognition pattern of CaM kinase I. They differ however in some respects, e.g. the replacement of Ser by Thr (accepted by AMP kinase but not by HRK-A nor SNF-1), the shape of the preferred hydrophobic chains at $n-5$ and $n+4$ and the displacement of the basic residue from $n-3$ to $n-4$, which is tolerated by AMP kinase and HRK-A but not by SNF-1 [32].

5.2.8. eIF2α kinases

The phosphorylation of Ser-51 of the α subunit of eIF2 is carried out by two protein kinases (HRI and PKR) sharing structural homologies with each other and a narrow substrate specificity. The site phosphorylated in eIF2α (ILLSELS^{51}RRRIR) and studies with peptide substrates support the concept that both these protein kinases recognize multiple basic residues on the C-terminal side of serine. The pattern which is most likely to represent a common consensus is SxRxxR [91]. Neither HRI nor PKR are expected to select a hydrophobic residue at position $n+1$ since the typical hydrophobic nature of the PKA p + 1 loop, LxxxPxxL (see Section 9.1.1), is altered in both, being replaced by SxxxTxxY and RxxxTxxY, respectively. It is possible that the two eIF2α kinases recognize different residues on the N-terminal side of serine, since the N-terminal extension of the peptide GSRRRRRRY to give RRRRYGSRRRRRRY decreases phosphorylation by HRI while increasing that by PKR [91]. Consistent with this, the residues homologous to PKA E170 (recognizing Arg at $n-2$, see Section 9.2) are Arg and Ser in HRI and PKR, respectively. It is possible therefore that Glu-49 at the phosphorylation site of eIFα (E^{49}LSR) is recognized by HRI but not by PKR. Moreover, from data obtained using clupeine as a substrate, PKR, but not HRI, seems to tolerate threonine instead of serine [91].

5.2.9. Pseudorabies and herpes simplex virus protein kinases

PRV-PK and HSV-PK are the only viral Ser/Thr protein kinases whose site specificity has been investigated in some detail [65,92]. These studies disclosed the importance of a cluster of N-terminal arginyl residues encompassing positions -5 to -2. Lysine could not replace arginine, while threonine was tolerated instead of serine. Apparently they are much more basophilic than PKA since the extension of RRASVA with two additional alanines to give RRRRASVA increases the phosphorylation efficiency by PRV-PK and PKA 126-fold and 7-fold, respectively [65]. It is also possible that these viral protein kinases do not share the predilection of most

basophilic protein kinases for a hydrophobic residue at position $n + 1$ since the crucial Leu at the end of the $p + 1$ loop motif of PKA (LxxxPxxL) is replaced by threonine, a feature also found in the Src family protein tyrosine kinases, thought to disfavour hydrophobic residues [12]. A proline at $n + 1$ is, in any case, a negative determinant [92]. An extended basic consensus sequence like that of these viral kinases is shared by the TGFβ receptor kinases (see Table 3) which, however, display a preference for lysine over arginine [14].

5.2.10. Pro-directed protein kinases

Cyclin-dependent kinases (cdc2 and various cdks) and mitogen-activated protein kinases (Erk1, 2 and 3, JNK 1 and 2 and SAPKs) constitute two growing families of protein kinases that share the absolute requirement for a prolyl residue at position $n + 1$ for substrate recognition and readily phosphorylate threonyl residues [33,106,159]. Threonine actually is quite frequent at their phosphoacceptor sites. Cyclin-dependent kinases differ from the mitogen-activated kinases because they do not display any marked preference for an additional prolyl residue at $n - 2$, which is a determinant for mitogen activated kinases (though not an absolute one, since it lacks in many physiological substrates). In contrast, basic residues downstream from proline $n + 1$, especially one at $n + 3$, are strongly selected by cyclin-dependent kinases, while they are not important in the case of MAP kinases. ERK-1 rather tends to prefer a third proline at position $n + 3$ (optimal sequence PxT/SPP in an oriented peptide library) (Ruzzene et al., unpublished). In the case of cyclin-dependent protein kinases, the nature of the associated cyclin should be also considered, as it tends to affect the specificity of the catalytic subunit [92a].

A ceramide-activated Ser/Thr protein kinase implicated in signal transduction initiated by TNF-α and I-1 [93] has also been shown to display proline-directed peptide specificity [94]. In contrast to cyclin-dependent and MAP-kinases, this kinase tolerates quite well a leucine residue interposed between threonine and proline, as peptides with the motifs PTLP and PLTLP are even better substrates than the reference peptide with the canonical motif of MAP kinases, PLTP. Moreover, while lysine residues adjacent to the N-terminal side of threonine are accepted by cdc2 kinase, they are disfavoured by the ceramide-activated kinase [94]. The effect of substituting threonine by serine was not checked.

5.2.11. Protein kinase CK2

Casein kinase-2 (or -II), recently renamed protein kinase CK2 to avoid confusion with the genuine casein kinase(s), is a pleiotropic and ubiquitous Ser/Thr protein kinase playing a central, albeit still enigmatic role in cell regulation [95]. Its relevance in a review devoted to the specificity of protein kinases is justified by the countless number of its substrates [96]. The extreme acidophilic nature of CK2, unequalled among protein kinases, with the possible exception of CK1, is somewhat unexpected considering its relatedness to Pro-directed and basophilic cyclin-dependent protein kinases. The specificity of CK2 has been thoroughly investigated, initially with the aid of casein and other artificial phosphoacceptor substrates (reviewed in [96]) and later by using a wide variety of peptide substrates [59,97–102] and by inspecting more than 100 natural phosphoacceptor sites [103]. Its minimum structural requirement is expressed by the consensus sequence S/TxxE/D; however multiple acidic residues besides the crucial one at the position $n + 3$ are required in order to obtain good phosphorylation efficiency and in fact clusters of acidic residues on the C-terminal side of serine/threonine are the hallmark of nearly all known CK2 sites. Often acidic residues are also present on the N-terminal side. The relevance of individual acidic residues at given positions appears to depend on the overall peptide structure: thus an acidic residue at $n + 2$ is quite important in peptides lacking acidic residues on the N-terminal side of serine [98], while its substitution by alanine in peptides having two acidic residues at $n - 1$ and $n - 2$ is without effect [104]. Conversely an acidic residue at $n + 1$ is almost as important as the crucial $(n + 3)$ one in peptides with acidic residues upstream [88] while being much more dispensable in peptides of the $S(E)_n$ series [98]. Likewise, interchangeability between D and E is greatly influenced by the structure, clusters of Asp being in general more beneficial than clusters of Glu, while a single substitution of Glu by Asp within a cluster of glutamic acid is detrimental at position $n + 3$ [103] but advantageous at $n + 2$ [104]. In general, wherever an acidic determinant is absent from

the sequence spanning position $n-2$ to $n+5$ the remaining ones become more important, giving rise, in some instances, to atypical sites, based on motifs Sp/D-Sp/D-*S*-Sp/D and phosphorylated despite the fact that they lack the canonical acidic determinant at $n+3$ [59].

The overall information emerging from these data is that the mode of peptide binding to CK2 is flexible and is determined by the combination of a number of interchangeable elements. This may partly account for the extraordinary number of CK2 phosphoacceptor sites and for the special relevance of compensatory effects, also observed with other protein kinases (Section 5). It is interesting to note in this context that CK2 can recognize as specificity determinants, besides Glu and Asp, phosphoserine and phosphotyrosine as well, this latter actually being the most potent determinant [105]. This puts CK2 in the category of 'phosphate-directed' protein kinases, similar to CK1, GSK3 and the Golgi apparatus casein kinase(s) (see below), with the notable difference however that CK2 is also *phosphotyrosine* directed and therefore potentially susceptible to substrate level modulation by protein tyrosine kinases. The outstanding acidophilicity of CK2 is also reflected by the detrimental effect of basic residues that are powerful negative determinants at any position tested between $n-1$ and $n+4$. A prolyl residue at $n+1$ is also a strong negative determinant [106], to a similar extent as a doublet of hydrophobic residues at $n+1/+2$ [103]. Serine is by far preferred over threonine as phosphoacceptor residue, while tyrosine is not phosphorylated at all, although it can bind to CK2 if it is surrounded by acidic specificity determinants, a feature that can be exploited for making peptide inhibitors.

5.2.12. Casein kinase (Golgi apparatus)

Bona fide casein kinases are typically enzymes concerned with the phosphorylation of casein in the lactating mammary gland. They are specifically located in the Golgi apparatus and a Golgi-enriched fraction casein kinase (GEF-CK) has been used for specificity studies with peptide substrates [30,101,104]. It is of interest to note that 'casein kinases' with the same or very similar specificity also appear to play a role in the secretory pathways of other tissues where they are responsible for the phosphorylation of SxEx motifs that are not suitable for CK2 [107]. The sequence of these Golgi casein kinases is currently unknown. Their specificity is dictated by the motif SxE/Sp, consistent with both the inspection of the sequences phosphorylated in casein fractions [108] and studies with peptide substrates [30,101,104]. These have provided evidence that GEF-CK does not tolerate the substitution of either Ser by Thr or Glu by Asp, within the triplet SXE. Only phosphoserine, but not phosphothreonine or phosphotyrosine, can replace glutamic acid as specificity determinant [104].

5.2.13. Protein kinase CK1

Like CK2, CK1 has recently been renamed to avoid confusion with the genuine casein kinases, as it was formerly termed casein kinase-1 (or -I). Despite the similar names and the common propensity to phosphorylate casein in vitro, CK1 is totally unrelated to CK2 in the protein kinase phylogenetic tree [25]. Moreover, whereas CK2 is a single entity, although with two alternative forms of the catalytic subunit (α and/or α'), CK1 is turning out to define a family of isoenzymes, seven of which are already known in mammals [109,110]. CK1 (termed at that time CK-S) provided the first example of a phosphate-directed protein kinase reaction, since it was shown that its phosphorylation sites in casein fractions were preceded by triplets of constitutively phosphorylated serines, whose dephosphorylation prevented subsequent phosphorylation by CK1 [111]. Such a phosphate-dependence was later confirmed with peptide substrates, showing that the crucial determinant is a single phosphoserine located at $n-3$ or, less effectively, at $n-4$ relative to the target serine [111a,112]. Phosphothreonine, but not phosphotyrosine, can replace phosphoserine as positive determinant [113]. Individual carboxylic residues are nearly ineffective. It has to be assumed, however, that multiple acidic residues on the N-terminal side of serine can replace phosphoserine as specificity determinants in certain circumstances, since protein phosphatase inhibitor-2 (I-2) [114] and DARPP-32 [115] are readily phosphorylated by CK1 at very acidic sites yet in the absence of any previous phosphorylation, and even peptides reproducing the CK1 sites in I-2 are phosphorylated with low K_m values [116]. The structural basis for such a high affinity, phos-

phate-independent, phosphorylation is poorly understood: while the phosphorylation of DARPP-32 may be driven by the large number of N-terminal acidic residues (a stretch of 18 consecutive E/D!), this hardly applies to the I-2 sites considering that the most acidic peptide (DDEEDEEM*S*ETADGE) is not phosphorylated as efficiently as a less acidic one (IGDDDDAY*S*DTETTEA). A possible explanation is that the presence of Asp instead of Glu at position $n-3$ is more important than the number of N-terminal acidic residues altogether. A hydrophobic residue at $n+1$ seems to play a beneficial role with CK1 [111,116].

Recently, two recombinant isoforms of CK1 from Saccharomyces pombe have been described as dual specificity kinases since they exhibited tyrosine autophosphorylation and activity on poly(E,Y)4:1 [40]. The Kcat with poly(E,Y)4:1 was 300-fold lower than that with casein, however. Using rat liver CK1 we were unable to find any detectable phosphorylation of the tyrosyl derivative of a 27-residue I-2 derived peptide whose seryl counterpart is an excellent substrate (Vaglio et al., unpublished). It is quite possible, on the other hand, that the numerous isoforms of CK1 display different substrate specificities. This would be also consistent with the observation that CK1δ undergoes unique regulation by an autophosphorylated autoinhibitory domain lacking the acidic character of CK1 sites and rather rich in prolines [117].

5.2.14. Glycogen synthase kinase-3

GSK-3 was firstly believed to be a highly specific glycogen synthase kinase, but subsequently several other substrates were detected for it, including inhibitor-2, the G subunit of PP1, c-Jun, Myb and c-Myc. Like CK1, GSK3 behaves as a phosphate-directed protein kinase, depending on a priming kinase for site recognition [118–120] and participating in well-documented hierarchical phosphorylations [121]. This feature was also studied with the aid of synthetic peptides, showing that the priming phosphoserine is located at $n+4$ [122]. However this recognition motif appears not to be the only one used by GSK3, since phosphorylation of some substrates (e.g. c-Jun, Myb and c-Myc) is seemingly not phosphate-directed and even two typically hierarchical sites in glycogen synthase (sites 3a and 3b) can be phosphorylated in mutants that lack the priming site-5, though the implication of another kinase in this case can not be ruled out either [123]. Whenever GSK3 is not phosphate-directed it does not seem to be Glu/Asp-directed either (as CK1 is) since acidic residues are poorly represented in its sites, which conversely are rich of prolines. The importance of proline for GSK3 would not come as a surprise considering its phylogenetic relatedness to the Pro-directed protein kinases of the cell division cycle.

5.2.15. G-protein coupled receptor kinases (GRKs)

The most studied members of this growing family of kinases [124] are rhodopsin kinase (RhK or GRK1) and the β-adrenergic receptor kinase (βARK or GRK-2). Despite the fact that they belong to the big 'AGC' group mainly composed of basophilic kinases (PKA, PKG, PKC, S6 kinases, etc.), βARK and RhK are generally assigned to the category of acidophilic kinases. It should be emphasized however that this assignment is based on data with peptide substrates, whereas GRKs are dedicated kinases with a narrow in vivo specificity confined to the activated (agonist-bound) receptor and not including an inactive or antagonist-bound receptor. Such targeting is dictated by micro-compartmentation mediated by binding to regions of the activated receptors that are not themselves actual substrates for the GRKs, but trigger the activation of the kinase [124]. Consequently, in vitro assays with peptide substrates GRKs display dramatically impaired catalytic efficiencies [125–127] and their specificity may be different from those in physiological conditions. Consistent with this caveat is the discrepancy between the information provided by peptide substrates (disclosing marked preference of βARK and RhK for peptides with acidic clusters upstream and downstream from serine, respectively) [128], and the residues affected by GRKs in rhodopsin (DDEAS^{334}TTVS^{338}KTETS^{344}QVAPA), which are those remote from acidic residues (S^{338}/S^{344}), while Ser-334, fulfilling the peptide consensus sequence, is not phosphorylated [124]. On the other hand, the sequence around the four serines phosphorylated by βARK in the α2A adrenergic receptor (DLEESSSSDHA) is more consistent with data with peptide substrates, but the finding that after mutating one to three serines the one(s) left are all equally phosphorylated [129] argues against a stringent requirement for

specificity determinants with definite spacing from the target aminoacid.

5.2.16. Putative kinase phosphorylating p70^{s6k} Thr-389

Recently Pearson et al. have identified a new phosphorylation site (Thr-389) within a conserved hydrophobic domain of p70^{s6k} which is the principal target of rapamycin-induced p70^{s6k} inactivation [130]. Since this site (situated downstream from the catalytic domain) does not reveal any obvious consensus for known protein kinases, it is postulated that it is the target for a kinase of novel specificity. It is interesting that the sequence motif around Thr-389 is conserved in several related kinases, its most peculiar features being two Phe at positions $n-4$ and $n-1$, a Gly at $n-2$, a Tyr or Phe at $n+1$ and a hydrophobic residue at $n+2$. If this sequence is important for phosphorylation, the kinase responsible would display a unique selection for hydrophobic determinants

5.2.17. 'Adoptive kinases' (PDHK, BCKDHK and DNA-PK)

Although the great majority of eukaryotic protein kinases are members of a single family of structurally related enzymes [25], a small number of enzymes have been described that phosphorylate proteins at Ser/Thr residues but are not related to this family [25]. Two are mitochondrial enzymes, one that catalyzes the phosphorylation of PDH and the other the phosphorylation of BCKDH. PDHK and BCKDK are structurally related to the so-called histidine kinases of procaryotes. They phosphorylate seryl sites containing acidic residues. In the case of PDH site-1, whose sequence is GHSMSDPG, the actual relevance of the aspartic acid at $n+3$ has been confirmed with peptide substrates in which this residue was replaced by Asn [131]. The motif SxxD is also found in two phosphoacceptor sites of BCKDH [132].

Data have recently accumulated showing that some of the PI3 kinase family members have protein kinase activity, a property that seemingly reflects the presence of conserved sequences between PI3 kinase and protein kinase families [133]. One of these, termed DNA-dependent protein kinase, which is involved in DNA double-stranded break repair, has been shown to phosphorylate proteins while having no detectable activity toward lipids [134]. Two motifs have been reported to be phosphorylated in vitro by DNA-PK, D/E/Q-S/T- [135] and P-S/T-X [136], emphasis being given to either the recurrent Q at $n+1$ or P at $n-1$, respectively. The actual relevance of these residues, however, was not checked by comparing substituted peptides.

5.3. Consensus sequences of individual tyrosine kinases

Reviewing the substrate specificity of PTKs, Gehalen and Harrison in 1990 [137] complained that studies of synthetic peptides 'have frustrated efforts to clearly define primary structure determinants'. 'It is perhaps of some consolation to investigators....that not all peptides containing tyrosine residues are substrates' they bitterly added. Their sense of frustration was quite understandable considering that most of the PTK peptide substrates derived from physiological substrates known at that time exhibited K_m values in the millimolar range, i.e. two to three orders of magnitude higher than the best peptide substrates of Ser/Thr kinases. Not surprisingly the scanty effectiveness of such substrates did not deteriorate much upon substitution of the residues suspected to act as 'specificity' determinants!

At about the same time, however, a spleen PTK termed TPKIIB was found to phosphorylate the Src-derived peptide EDNEYTA with a K_m value of 58 μM, and, more important, the K_m increased 30-fold (with a concomitant drop of Kcat to negligible values [138] if the acidic residues upstream from tyrosine were substituted by neutral ones. This behaviour was reminiscent of that of highly 'site specific' Ser/Thr protein kinases. Today peptide substrates displaying K_m values in the low micromolar range with many PTKs are known (see Table 4) and the dogma that these kinases invariably exhibit low affinity toward peptide substrates has been dismissed.

5.3.1. Syk

TPKIIB has recently been shown to be identical to the catalytic domain of p72^{syk}, belonging to a small family of non-receptor PTKs (including also p70^{zap}) implicated in the activation of hematopoietic cells. p72^{syk} itself has been shown to be endowed with a remarkable site specificity, exhibiting K_m values as

Table 4
Peptides phosphorylated with high affinity by protein tyrosine kinases.

PTK	Peptide	K_m (μM)	Ref.
Syk	ENE$\underline{Y}$TA	58	[138]
	EQEDEPEGD$\underline{Y}$EEVLE	4	[60]
c-Fgr	EDENL$\underline{Y}$EGLNLDDCSM$\underline{Y}$EDI	33	[60]
c-Src	YI$\underline{Y}$GSFK	55	[15]
	AEEEI$\underline{Y}$GEFEAKKKK	33	[12]
Lyn	c(EDNE$\underline{Y}$TA)$_2$	20	[181]
Abl	EAI$\underline{Y}$AAPFAKKK	4	[12]
CSK	EEEPQFEEIP$\underline{Y}$LELLP	40	[a]
I-R	KKSRGD$\underline{Y}$MTMQIG	30	[150]
IGF1-R	KKKSPGE$\underline{Y}$VNIEFG	26	[150]

[a] Ruzzene et al., unpublished.

low as 4 μM with peptides derived from HS1 and Vav proteins, and a stringent requirement for acidic residues on the N-terminal side of tyrosine [60]. The position at which an individual acidic determinant is most effective is $n-1$; however, an acidic residue at $n-1$ is dispensable if acidic residues are present at positions -3, -4, +1 and +2. Acidic residues at $n+1/+2$ play a favourable role even if acidic residues are already present on the N-terminal side, while at position $n+3$ a hydrophobic residue is preferred, a circumstance that enables $p72^{syk}$ to generate binding sites for the SH2 domains of the Src family PTKs. It is remarkable that $p72^{syk}$ tolerates the replacement of the acidic determinants by phosphorylated residues of tyrosine and, even better, threonine [139,60], rendering it the only 'phosphate-directed' PTK known to date.

5.3.2. Src and Abl PTKs

Recently good peptide substrates have also been obtained for other non-receptor PTKs, including several members of the Src family and Abl (see Table 4). A common denominator of these kinases, and also of the members of the Fps/Fes and CSK families, appears to be the requirement for a hydrophobic residue at $n-1$ for optimal phosphorylation. The relevance of this specificity determinant has been amply confirmed using either conventional peptides [60] or variably degenerate peptide libraries [9,12,13], and it can be considered the hallmark that distinguishes these enzymes from Syk as well as from the majority of receptor PTKs [12].

The additional specificity determinants of Src and Abl are less clearly understood: the ability to recognize acidic residues on the N-terminal side of the target residue, suggested by the structure of many phosphotyrosyl sites, is only partially consistent with the information provided by an oriented peptide library, whose data support the view that while v-Src is modestly acidophilic, Lck weakly selects a single Glu at $n-2$ and Abl does not select acidic residues at all [48]. The outcome of a degenerate library of heptapeptides in which the position of tyrosine is not pre-determined is even more drastic, supporting the view that the residues upstream from position $n-2$ are of no importance, since the peptide selected by c-Src is YI$\underline{Y}$GSFK, entirely devoid of acidic residues [13]. It may be worth noting that the good K_m of this peptide is emulated by a quite unrelated peptide, AEEIYGEFEAKKKK (see Table 4), their only common feature being Ile at $n-1$. This suggests that, apart from the hydrophobic/β-branched side chain at $n-1$, there are no individual determinants stringently required at any other given position, but, rather, there may be certain overall peptide structures that are better suited than others to fit in the catalytic site of Src kinases. The modest relevance of acidic residues for Src-mediated phosphorylation is also supported by the similar phosphorylation of poly(E_4Y) and poly(K_4Y) [140], by the detection of protein substrates lacking acidic residues around phosphotyrosine [12,49,50,137] and by the small effect of substituting acidic residues in peptide substrates [138,141]. It should also be considered, however, that the specificity determinants of the different members of the Src family may vary significantly, as suggested by the results of studies with an oriented peptide library comparing Src and Lck [12] and comparative studies on Lyn and c-Fgr [138,142]. Unlike c-Fgr, Lyn appears to dislike acidic residues [143,144], at least when it is assayed in the absence of polylysine, which is not in agreement with the results of studies with an oriented peptide library (see Table 3). c-Fgr, on the other hand exhibits an astonishing promiscuity, enabling it to appreciably phosphorylate, though with high K_m values, tiny di- and tri-peptides like AAY, GAY and AY (see Ref. [142], where c-Fgr is termed TPK-III). This unusual feature of c-Fgr has been exploited for the preparation of a large variety of phosphorylated peptides for use in studying the

specificity of PTPases [145]. Pro at $n+1$ has been found to prevent phosphorylation by c-Src [140]. It is also negatively selected in the peptides phosphorylated within a peptide library, suggesting that it may play a negative role in tyrosine phosphorylation by a variety of PTKs.

5.3.3. CSK

C-terminal Src kinase (CSK) specifically phosphorylates Src family members at their C-terminal tyrosine (equivalent to Y-527 in c-Src) and down-regulates their activity. CSK is poorly active on peptides reproducing this site, suggesting that its specificity is dictated by higher order structural features [146]. It came as a surprise, therefore, that CSK can efficiently phosphorylate a middle-T antigen-derived peptide only distantly related to the C-terminal site of Src kinases (see Table 4). In this peptide the tyrosine preferentially affected by CSK (Tyr-330) is embedded in the middle of an extraordinarily hydrophobic sequence, suggesting that this feature is responsible for the good performance of CSK. This conclusion has been corroborated with an oriented peptide library (see Table 3).

5.3.4. Receptor PTKs

Receptor PTKs, once activated by the extracellular agonist, typically become autophosphorylated at various regions of their intracellular domains, thus creating docking sites for cytoplasmic proteins that bind to them through their SH2 domains. It is doubtful if autophosphorylation or even phosphorylation of proteins associated with the receptor PTKs through their SH2 domains is strictly dependent on local specificity determinants that are recognized by the active site of the kinase. However many specificity studies on receptor PTKs have been performed using peptides that reproduce these autophosphorylation sites [147,148]. In some cases peptides derived from the main substrate of insulin receptor (IRS-1) have been used [149,150]. Recently an oriented partially degenerate peptide library has been employed to investigate the specificity of several protein kinases, including four receptor ones, EGFR, PDGFR, FGFR and I-R [12]. A generalization that can be drawn from these studies is that receptor PTKs do not share the requirement of most non-receptor PTKs for a hydrophobic residue at $n-1$, where they prefer an acidic residue [12,147,150], similar to the cytoplasmic PTK, Syk. Weak to medium selectivity of glutamic acid at positions $n-2$ and $n-3$ is also found with the peptide library. In contrast the residues selected on the C-terminal side of tyrosine are mostly hydrophobic, often including methionine, which appears to play an especially important role with the insulin receptor PTK [149]. This kind of selection provides optimal sites for type III SH2 domain binding. Data from both peptide libraries [12] and conventional peptides [148–150] support the view that changes in specificity among receptor PTKs are mainly dictated by the nature of the residues C-terminal to tyrosine.

6. 'Hierarchical' protein phosphorylations

As outlined in the previous chapter there are several protein kinases that recognize previously phosphorylated residues as specificity determinants: these include Ser/Thr protein kinases, namely GSK3, CK1, CK2 and the 'casein kinases' of secretory pathways, and also at least one tyrosine kinase, Syk. Whenever this is the case the occurrence of a 'primary' phosphorylation will generate a phosphoacceptor site suitable for a 'secondary' phosphorylation, which can be followed by another, as was elegantly shown by Roach and colleagues [122] for the phosphorylation of the glycogen synthase sites 3a, 3b, 3c and 4 by GSK3, which is triggered by the priming phosphorylation of site-5 by another kinase, CK2. In this as well as in other examples the 'primary' and 'secondary' agents are distinct kinases. It is also possible, however, that the same kinase catalyzes both the primary phosphorylation and the secondary one(s), as in the case of multiple phosphorylations by CK2 of DARPP-32 and c-Myc. Roach coined the term 'hyerarchical phosphorylation' to indicate this sort of orderly interdependent multiple phosphorylations taking place at contiguous sites and involving phosphorylatable residues with definite spacing between each other [121]. One interesting possibility raised by the occurrence of hierarchical phosphorylation is for 'cross-talk' between Ser/Thr kinases and Tyr kinases. The potential occurrence of this reciprocal influence is suggested by the observation that, on one hand Tyr-P can act as a powerful recognition determinant of the Ser/Thr kinase CK2 [105], and, on the

other, a tyrosine kinase, p72^{syk}, recognizes as specificity determinants phosphorylated threonine and, to a lesser extent, serine [139].

Recently examples of interdependent multiple tyrosine phosphorylation have been provided that are reminiscent of canonical hierarchical phosphorylation, but they differ in that the primary phosphorylation occurs at regions of the protein substrate that are remote from the secondary phosphorylation site(s) and are recognized by a Src homology-2 (SH2) domain of the secondary kinase rather than by its catalytic domain [49,165–168]. The primary and secondary agents can be either the same PTK [49,165,166] or distinct enzymes [167,168]. To avoid confusion with the canonical mode of hierarchical phosphorylation (which could be specified as 'homotopic', to indicate that primary and secondary phosphorylations occur in the same region) it may be advisable to use different terms to indicate the SH2 mediated 'allotopic' hierarchical phosphorylations. The name 'processive' phosphorylation has been proposed in the case of a single kinase (Abl) being responsible for both the primary and secondary phosphorylations [165,166] while the allotopic hierarchical phosphorylation of HS1 protein carried about by p72^{syk} and Src kinases has been termed 'sequential' phosphorylation [168].

7. A quick look at protein phosphatases

As protein phosphorylation is a reversible process, the residues phosphorylated by protein kinases are expected to undergo well-timed dephosphorylation by protein phosphatases. This, of course, implies tight co-ordination between kinases and phosphatases. Pertinent to this may be the question as to whether the specificity determinants of protein kinases are also implicated in the dephosphorylation process. The problem may be even more complicated in the case of phosphotyrosyl sites, where the elements for recognition by SH2 and possibly PTB domains coexist with those recognized by the kinases and the phosphatases.

As far as we can say protein phosphatases do not recognize sensu stricto consensus sequences like those of many kinases; undoubtedly, however, dephosphorylation is greatly influenced by a number of local structural features, some of which also appear important for phosphorylation (reviewed in [169]). Some of them that may deserve special attention are summarized below.

i. Ser/Thr protein phosphatases, especially the most pleiotropic species, PP2A, and with the notable exception of calcineurin (PP2B) [170] display a striking preference for phosphothreonyl residues over phosphoseryl residues included within identical sequences [171–173]. Curiously, the majority, albeit not all, of protein kinases exhibit an opposite preference. Consequently, phosphoseryl sites are more far numerous that phosphothreonyl ones, although the abundance of threonine in proteins is on average quite similar to that of serine. This could reflect a mechanism to prolonge the life of phosphorylated residues protecting them from premature dephosphorylarion.

ii. A proline at $n + 1$, which is disliked by the majority of protein kinases but is absolutely required by a subset of Pro-directed kinases, is a strong negative determinant for protein phosphatases, with the exception again of calcineurin (PP2B). Whenever the two detrimental features, i.e. phosphoserine and Pro at $n + 1$, are combined they generate sites that are refractory to PP2A, while dephosphorylation can be rescued if SerP is replaced by ThrP [174]. This could partly account for the relatively high frequency of TP as opposed to SP motifs at the sites phosphorylated by Pro-directed protein kinases, while seryl sites predominate much more with the other kinases. It is also possible that the phosphorylated motif SpP is an indicator of stable phosphorylation and/or of calcineurin targeting.

iii. Basic residues on the N-terminal side of the phosphorylatable aminoacid are recognized as positive determinants by the majority of Ser/Thr protein kinases. Multiple N-terminal basic residues have also been shown to facilitate the otherwise slow dephosphorylation of phosphoseryl residues by protein phosphatases 1, 2A and 2B [170]. In order to observe this latter effect, however, more basic residues are necessary than are normally needed by most protein kinases. It has been proposed that such a redundancy of basic residues, evident, e.g., in phosphorylase (see Section 5.2.6), is required for the dephosphorylation of otherwise refractory sites [175].

iv. Acidic residues on the C-terminal side of the

phosphoaminoacid are well tolerated by PP2A [172], but not by calcineurin [170] or by PP2C [173]. Consistent with this, PP2A appears to be primarily responsible for the dephosphorylation of the highly acidic site phosphorylated by CK2 in its numerous substrates. Examples of this correlation are provided in Refs. [95,176,177].

v. As discussed elsewhere [169], efficient dephosphorylation of phosphotyrosyl substrates is critically dependent on the presence of acidic residues located at varying distances on the N-terminal side of TyrP. These acidic residues are actually a recurrent feature of most tyrosine phosphoacceptor sites, a fact often interpreted as suggesting an acidic requirement of the kinases. As discussed above, many protein tyrosine kinases do not really need or just weakly select acidic determinants N-terminal of tyrosine. This supports the feeling that these acidic residues are rather the hallmark of faster dephosphorylation.

8. Importance of higher-order structures

A vexed question about the phosphoacceptor sites is whether their 3D structure could be more important than their lineal sequence in determining their phosphorylation. Early studies with protein kinases disclosed a significant correlation between susceptibility to phosphorylation and inclusion into sequences predicted to assume β-turn conformation [178,179]. Subsequent studies with peptide substrates for both Ser/Thr [101] and Tyr protein kinases [138,180] supported the concept that, lacking stronger determinants, a β-turn or any exposed and flexible structure could indeed represent a feature that facilitated phosphorylation. While it is conceivable and in a sense obvious that the presentation of phosphoacceptor sites to the catalytic domain of protein kinases may be assisted by an exposed and/or a bent conformation like a β-turn or a loop, it is now clear from the crystal structure of the PKA-PKI peptide binary complex that the consensus sequence surrounding the phosphoacceptor residue is accommodated in an extended conformation inside a cleft between the two lobes of the kinase [67]. This would imply that the previous structure of the peptide, if any, is altered upon binding, and suggests that flexibility of the sequence around the phosphoacceptor residue will be a favourable feature. Consistent with this, the modest phosphorylation efficiency by Lyn of a cyclic heptapeptide having a rigid structure could be improved more than 10-fold merely by dimerization, giving rise to a more disordered structure whose lineal sequence around tyrosine was unchanged [181].

The answer to another question, i.e. how can the tertiary structure of the protein substrate influence phosphorylation, is less straightforward and partially outside the scope of this review, which mainly focuses on local specificity determinants. There are two aspects of the problem that may deserve a short comment, however. The first is that the tertiary structure can have an overriding influence and prevent the phosphorylation of otherwise suitable sites simply because they are hidden inside the folded molecule. The less trivial possibility also exists, however, that in some cases the proper proximity of the essential specificity determinants and the phosphoacceptor residue is defined by the folding of the substrate in its native conformation. In such a case a 'tridimensional' instead of lineal consensus sequence would be generated, whose persistence will crucially depend on the native conformation of the protein substrate. That this is not mere speculation is suggested by a number of observations. An example is provided by the hierarchical phosphorylation of I-2 by GSK3 at Thr-72. Normally GSK3 recognizes the consensus sequence SxxxSp (see Table 3). In this case, however, the phosphate acting as specificity determinant is that of Ser-86, 14 residues downstream [182] whose correct positioning is conceivable only assuming a loop between Thr-72 and Ser-87. Other examples of how regions remote from the phosphoacceptor site are required for efficient phosphorylation are provided by various substrates of CK2, a kinase that otherwise exhibits a canonical, 'lineal' site specificity. Thus the phosphorylation of Ser-2 of the β-subunit of CK2 is dependent on a cluster of acidic aminoacids more than 50 residues downstream [183] and the phosphorylation by CK2 of the penultimate residue of p53, a serine that does not fulfil the primary structure requirement of this kinase, is drastically reduced in a truncated p53 expression mutant that lacks the N-terminal domain [103]. Even more striking is the phosphorylation of Ser-56 of the mannose-6-P receptor lying in the sequence EESEERD with a detrimental R at the crucial $n+3$ position: the efficient

phosphorylation of this site requires the upstream 26-32 sequence ADGCDFV containing two aspartic acids [176]. These data imply higher order structure requirements for generating phosphoacceptor sites normally determined by lineal consensus sequences. The crystal structure of NDPK [184] is very telling in this respect as it shows that a number of acidic residues quite remote from Ser-122 (and Ser-125) are brought together by the folding of the molecule in the vicinity of these residues, which are phosphorylated by CK2 despite the fact that they do not conform to the lineal consensus SxxE/D [185].

9. The structural bases for kinase selectivity

'A review of protein kinase recognition sequence...is premature because we do not know the three-dimensional structure of a single protein kinase substrate complex' noted Kemp and Pearson in 1990

Table 5
Residues of protein kinases that have been shown to interact with peptide substrates/inhibitors, and their homologs in PKA

Kinase	PKA [a]	Position recognized [b]	Ref.
PKA:			
E170, E130	–	$n-2$	[67,193]
E127, E331, T51	–	$n-3$	[67,187,193]
L198, P202, L205	–	$n+1$	[67]
E203	–	$n-6/-3$	[66,67]
Y^{235}PPFF	–	$n-11$	[66,67]
Ph. kinase:			
E154	E170	$n-2$	[87]
E111	E127	$n-3$	[87]
MLCK (sm.m.):			
E777	E127	$n-3$	[196,197]
E821	E170	$n-3$	[196,197]
MLCK (sk.m.):			
E377	E127	$n-3$	[196,197]
E421	E170	$n-2$	[196,197]
Twitchin:			
E6,067	E170	K6,224 [c]	[190]
E6,023	E127	R6,237 [c]	[190]
E6,026	S130	K6,222 [c]	[190]
E6,112	L211	R6,255 [c]	[190]
K6,097	W196	D6,256 [c]	[190]
RhK:			
K491	E333	N-terminal	[204]
CK2:			
H160	E170	$n-2$	[198]
R191,R195,K198	L198,P202,L205	$n+1$	[88,199]
K79(R80)	K83(Q84) [d]	$n+3$	[88]
K^{74}KKK	K^{78}VVK [d]	$n+4/+5(+3)$	[88]
CK1:			
R183, K229,G215	E203,Q241,A239	$n-3/-4(PO_4^{2-})$	[191,202]
K^{217}AATKR	[e]	N-terminal	[202]

[a] PKA residues homologous to residues shown to be implicated in substrate recognition by other protein kinases.
[b] Expressed relative to the phosphoacceptor residue or to the corresponding alanine in the PKI peptide (position 'zero').
[c] Residues of the carboxyl-terminal tail interacting with the indicated residues of the catalytic domain of twitchin. The identification of position 'zero' in this case is impossible.
[d] According to the manual alignment between CK2 and PKA adopted in [88].
[e] A series of gaps have been introduced in PKA in this region to allow best alignment with CK1 [202].

[1]. One year later the efforts of the Taylor laboratory provided a three-dimensional structure of the catalytic subunit of PKA as a binary complex with a pseudosubstrate peptide inhibitor [67,186]. From these and the studies of others [187] the modes of interaction of both the phosphoacceptor substrate and ATP and the likely mechanism of catalysis have been elucidated. These data in conjunction with the resolution of the crystal structure of other protein kinases revealing a common architectural motif [28,188–191] provided the structural rationale accounting for the site specificity of these enzymes and enormously facilitated the design of mutational experiments aimed at defining the structural elements determining the consensus sequences of other protein kinases as well.

The most evident common feature of protein kinases is a bilobal structure with a deep cleft between the small (upper) lobe and the large (lower) one. The deeper portion of the cleft is filled by ATP-Mg, presumably in close contact with the phosphoacceptor residue of the peptide substrate, while the surrounding residues of the peptide are wedged in the rest of the cleft in an extended conformation. In the case of PKA the majority of the peptide binding sites are located on the large (lower) lobe. This may be a common feature of protein kinases that recognize specificity determinants located on the N-terminal side of Ser/Thr and at position $n + 1$, as in the case of PKA. It is possible however that whenever the crucial specificity determinants are located downstream from position $n + 1$, as in the case of CK2, the upper small lobe may play a predominant role in substrate recognition (Section 9.1.3).

9.1. Residues implicated in substrate recognition

The residues which have been shown or suggested by crystallographic and/or mutational studies to be implicated in substrate recognition by various kinases and their counterparts in the peptide substrates are listed in Table 5. It should be recalled that no crystal structure of a bona fide phosphoacceptor substrate bound to a kinase has yet been solved. However, one would expect that this is most closely mimicked by the crystals of PKA complexed with a 20 aminoacid peptide inhibitor [67], whose specificity determinants in its C terminal segment (14–20) have been shown to coincide with those of a peptide substrate in which Ala-17 was replaced by serine [66]. PKA has therefore been taken as the reference kinase in Table 5 and the residues of PKA which are homologous to residues of other kinases implicated in substrate recognition have also been indicated whenever alignment is possible.

9.1.1. PKA

The acidic residues of PKA that contribute to peptide recognition by electrostatic interactions with the crucial basic side chains of arginine at positions $n - 2$ and $n - 3$ were identified by chemical modification [192] and charged-to-alanine scanning mutagenesis analysis of yeast PKA [193,194]. Using the numbering of mammalian PKA adopted by Taylor and coworkers, which is followed in Table 5, E-170 and E-230 were found to interact with position $n - 2$, whereas E-127 is the counterpart of the arginine at $n - 3$. The analysis of the crystal structure of the PKA catalytic subunit complexed with the peptide inhibitor [67] confirmed the occurrence of these interactions and also showed an interaction of R at the $n - 3$ position in the inhibitory peptide with E-331, situated outside the catalytic domain, in the C-terminal part of the molecule. In the ternary complex with an ATP analogue [187] this latter interaction is not present: in that model the side chain of Arg at $n - 3$, besides making an ion-pair with E-127, forms a hydrogen bond with Thr-51, inside the 'Gly-rich loop'. The crystallographic studies also accounted for the preference for a hydrophobic residue at $n + 1$ by showing that L-198, P-202 and L-205 provide a hydrophobic pocket (the '$p + 1$ loop') to harbour this residue [67]. The high affinity binding of the inhibitor peptide is ensured by additional interactions besides the ones provided by the consensus sequence (-3, -2 and $+1$). An arginine at $n - 6$ and a phenylalanine at $n - 11$ are especially important in this respect [66,67]. The structural analysis showed that R at $n - 6$ interacts among others with E-203, while a hydrophobic pocket formed by residues 235–239 accommodates F at $n - 11$ in the inhibitor peptide. Interestingly, however, if the inhibitor peptide is converted into a substrate by changing Ala at zero position to Ser, both R at $n - 6$ and F at $n - 11$ loose their importance [66], suggesting that the mode of binding of the pseudosubstrate inhibitor and of the genuine substrate are partly different. Incidentally,

this also raises the possibility that other modifications of the residue at position zero, especially substitution of serine by threonine, might significantly influence the mechanism of substrate binding. Pertinent to this is the possibility that the efficient phosphorylation of threonyl targets (e.g. protein phosphatase inhibitor-1) which are intrinsically poor substrates of PKA (as well as of many other protein kinases) is improved by relatively remote structural features that are not recognized in seryl substrates.

Recently it has been shown that His-87, playing a unique role in formation of a contact site between the two lobes by interacting with the essential activating autophosphorylation site Thr-197, is also implicated in the recognition of residues at the $n + 2$ position [195]. Substrate positions downstream from $n + 1$ are normally unimportant for PKA recognition; the interactions with the pseudosubstrate regions of PKI and the regulatory subunits may represent an exception however (see Section 9.1.3).

9.1.2. Calmodulin-dependent protein kinases

Mutational studies with phosphorylase kinase [87] and MLCK from either smooth muscle or skeletal muscle [196,197] have confirmed the implied role of acidic residues homologous to PKA E-170 and E-127 in the recognition of basic residues situated at positions $n - 2$ and $n - 3$ in peptide/protein substrates, respectively. Interestingly, the mutation of a number of additional acidic residues of MLCK implicated in binding to the autoinhibitory domain only marginally affects the phosphorylation of the substrate myosin light chain [58] supporting the concept (see Section 9.1.1) that protein kinases can bind substrates and pseudosubstrate (auto)inhibitory sequences through distinct pathways.

This point of view is corroborated by the crystal structure of the twitching catalytic domain complexed with 60 residues of carboxyl-terminal tail responsible for intrasteric down-regulation [190]. This C-terminal peptide extends through the active site making extensive contacts with the catalytic core. Indeed it mirrors the active site and is more than a simple pseudo-substrate. Two interactions are mediated by E-6067 and E-6023 which are homologous to PKA E-170 and E-127, respectively, that recognize basic residues at positions $n - 2$ and $n - 3$ (see Section 9.1.1); however twitching E-6067 and E-6023 do not interact with adjacent basic residues as do their PKA counterparts, but with residues which are far apart from each other in the primary structure of the down-regulatory domain, namely K-6224 and R-6237.

Taken together, the data from calmodulin-dependent protein kinases reinforce the conclusions drawn from PKA, by showing that the residues homologue to PKA E-170 and E-127 play a general role in the recognition of specificity determinants at $n - 2$ and $n - 3$, respectively and that additional residues, which are not so important in normal substrate recognition are implicated in the high affinity binding of inhibitory peptides/domains.

9.1.3. CK2

Unlike PKA and the calmodulin-dependent protein kinases, CK2 is acidophilic. It is quite telling therefore that some paradigms of substrate recognition by basophilic protein kinases have also been confirmed with CK2, reinforcing the view that homologous residues are utilized for substrate recognition by kinases exhibiting different specificities. The CK2 homologs of PKA E-170 and L198/P202/L205 (the so called '$p + 1$ loop') are H-160, and R-191/R195/K198, respectively. By mutational analysis it was shown that they play the same role as in PKA, i.e. they recognize an acidic residue at positions $n - 2$ (in peptides where this residue is important) and $n + 1$, respectively [198,199]. Mutational analysis of CK2 also disclosed one or two basic residues (K79/R80) that interact with the most crucial specificity determinant at position $n + 3$, and a cluster of four lysines (K74-K77) whose mutation impairs recognition of acidic determinants at $n + 4/ + 5$ and, to a lesser extent, $n + 3$ [88]. These basic residues are located at the border between subdomains II and III, where a series of several gaps has been introduced in most protein kinases to allow best alignment. By manual alignment with PKA the 74-77 basic stretch of CK2 falls into the end of subdomain II rather than at the beginning of subdomain III [88] and in this alignment K79 becomes homologous to PKA K83, whose side chain faces an aspartic acid at position $n + 3$ in the PKI peptide co-crystallized with it [67]. Although it has never been reported that PKA recognizes acidic residues in its substrates at positions downstream from serine it is noteworthy that in all the natural inhibitory polypeptides of PKA, PKI,

RI and RII, an acidic residue is present at position $n+3$ or $n+4$ of their pseudosubstrate sites [200]. Pertinent to this may be also the observation that PKG, in which the basic nature of PKA K83 is conserved, appears to select for sequences with Asp at $n+3$ [201].

9.1.4. CK1

Among Ser/Thr protein kinases whose crystal structure has been solved CK1 is the one most distantly related to PKA. It is remarkable therefore that, despite the low similarity in primary structure, the overall bilobal architecture of PKA is conserved in it, albeit with a number of unique features. CK1 was crystallised in active conformation without any substrate/inhibitor bound to it. However, individual sulphate [191] and tungstate [202] anions bound to the enzyme in the same place are likely to map a potential primary point of phosphorylated substrate recognition. To be effective this phosphate group has to be located at $n-3$ or, less efficiently $n-4$ position in the peptide substrate (see Table 3)). The residues interacting with sulphate and tungstate are R183, K229 and G220, homologous to PKA E-203, Q241 and A239 (see Table 6). E203 in PKA interacts with Arg at position $n-6$ of the inhibitor peptide [66,67]. The key residues involved in $n-2/-3$ substrate binding in PKA, E-127, E-170 and E-230 are replaced in CK1 by S-91, N135 and Y210, which are not suited for electrostatic interactions with negatively charged groups. It should be mentioned, however, that all known forms of CK1 are positively charged at neutral pH and the positive charges are enriched in the putative substrate binding region [191]. In particular a disordered loop which is adjacent to the tungstate-binding site in mammalian CK1δ

Table 6
Residues of representative kinases homologous to those that have been shown to interact with phosphoacceptor substrates

Position(s) in substrate:	$n-3$ [a]	$n-2$ [a]		$n+1$ [a]	$n+3/+4/+5$ [a]
PKA homologs	127 [b]	170 [b]	230 [b]	205 [b]	78–84 [b]
PKA	**E**	**E**	**E**	**L**	KVVKLKQ
PKC	D	D	E	I	VVIQDDD
p90^{s6k} (N)	D	E	T	M	TLKVRDR
CaM kin.II	E	E	I	L	KKLSARD
MLCK	**E**	**E**	I	V	SAKEKEN
PhK	**E**	**E**	T	L	GGGSFSAEE
cdc2	D	Q	E	R	DESEGVP
ERK	D	S/A	E	R	EHQTYCQ
GSK3	T	Q	E	R	KRFKNR–
CK2	D	**H**	S	**K**	**KKKKIKR**
βARK	D	A	K	M	RIKMKQG
CK1	S	D	Y	M/A [c] or L [d]	[c]
c-Src	S	A	E	T	TMSPEAF
Lyn	S	A	E	T	TMSVQAF
c-Fgr	S	A	E	T	TMSPKAF
c-Abl	N	R	E	T	TMEVEEF
p72^{syk}	S	R	E	Y	ANDP[ALK]DEL
CSK	S	R	E	T	ATAQAFL
Insulin R.	D	R	E	M	ASLRERI
PDGF R	D	R	E	M	ARSSEQA
EGF R	C	R	E	M	SPKANKE

Bold type denotes references residues that have been actually shown to be implicated in substrate recognition (see Table 5).
[a] Relative to the target aminoacid.
[b] PKA numbering according to Taylor and associates [67]. Alignment with CK2 in the 78–84 stretch was done manually [88].
[c] Variable in CK1 isoforms.
[d] Depending on the alignment adopted (see Section 9.1.4).

(K^{217}AATKRQ) has been proposed to be responsible for the binding of the anionic but not phosphorylated substrates [202]. This stretch has no homologous sequence in PKA. The concept that CK1 may use a *sui generis* pattern for substrate recognition is also suggested by the finding that the hydrophobicity of the loop p + 1 is not conserved in it (if the alignment in [202] is adopted) despite the fact that a hydrophobic residue appears to be preferred at position $n + 1$. By adopting the alignment used in [203], however, the crucial residue at the base of the pocket is Leu, consistent with the selection of a hydrophobic residue at $n + 1$.

9.2. Prediction of site specificity from primary structure

An interesting outcome of the studies summarized in Table 5 is that a number of homologous residues in different protein kinases appear to play a similar role in the recognition of specificity determinants placed at definite positions of the phosphoacceptor substrate. The general validity of this concept, at least in the case of Ser/Thr kinases and with possible exceptions (notably CK1), is corroborated by examining the nature of these key residues in Ser/Thr kinases not yet analyzed by mutational/crystallographic studies, but whose specificity determinants are already known. Several representative kinases are listed in Table 6 where their residues that are homologous to aminoacids shown to participate in the recognition of specificity determinants at definite position by either PKA [67,187,193] or CK2 [88,198,199] are indicated. The residues considered are those that in PKA interact with the basic determinants at $n - 2$ and $n - 3$ and with the hydrophobic determinant at $n + 1$; in this latter case only the third residue of the triplet which is located at the base of the pocket and appears to have the main contact with the side-chain at $n + 1$ [11], has been considered. This has actually been shown to play the major role in the recognition of an acidic residue at $n + 1$ by CK2 (Sarno et al., unpublished). Table 6 also shows the sequences that overlap by manual alignment the CK2 K74-R78 basic cluster which has been shown to interact with C-terminal determinants at positions $n + 3$, $+4$ and $+5$ [88].

Taken together the data of Table 6 are consistent with the idea that some prominent features of many Ser/Thr protein kinases can be predicted from the analysis of a number of signatures at fixed positions along their primary structure. As already outlined (e.g. Refs. [103,193] the nature of the residues homologous to PKA E170, E230, and E127 is especially indicative of the preference for either basic or acidic residues on the N-terminal side of the target aminoacid. In particular the acidic nature of E170 is conserved in all the protein kinases that recognize crucial basic determinants on the N-terminal side of serine, while it is lost in the other kinases (with the exception of CK1). Interestingly, moreover, two acidophilic kinases known to recognize carboxylic side chains close to the N-terminal side of serine, CK2 and βARK, have a basic residue instead of E-170 and E230, respectively, whereas basic residues at these positions are invariably absent in all the other kinases. Another quite diagnostic feature appears to be the nature of the residue homologous to PKA L205, at the base of the p + 1 hydrophobic pocket: a hydrophobic residue reflects selection of a hydrophobic residue at $n + 1$, whereas lysine correlates perfectly with the recognition of an acidic residue, as in the case of CK2, where this complementarity has been also demonstrated by mutagenesis [88,199]. Interestingly, an arginine at this position seems not to match an acidic residue, but rather a prolyl residue, as also suggested by space-filling models [11]. Actually an Arg replacing PKA L205 is a hallmark of all Pro-directed kinases, either cyclin- or mitogen-activated, and is also found in GSK3 which can be considered a *sui generis* Pro-directed kinase (Section 5.2.14).

The analysis of the residues equivalent to the CK2 basic stretch that recognizes residues downstream from $n + 2$ is not unequivocal due to the presence of a series of gaps permitting alternative alignments. Following the manual alignment used in [88], assigning homology between CK2 K79 and PKA K83, both in close contact with the substrate residue at $n + 3$, it becomes apparent that either acidic or basic residues predominate in protein kinases known to recognize C-terminal basic (PKC, cdc2) and acidic residues (GSK3, CK2, βARK), respectively. In MAP kinases, which are similar to cdc2 in being Pro-directed but differ in lacking basophilic nature (see Section

5.2.10), this segment is devoid of the acidic character found in cdc2.

In contrast to most Ser/Thr protein kinases (with the exception of CK1), Tyr kinases do not appear to conform to the pattern of predictions just outlined. In fact, as shown in Table 6, the nature of the postulated key residues in PTKs hardly reflects their differences in site specificity. It should be remembered in this connection that consensus sequences of PTKs generally are less stringent than those of Ser/Thr kinases and that one of the key residues (homologous to PKA E230) is invariably a glutamic acid in all PTKs despite the fact that these enzymes are believed to display a more or less pronounced preference for acidic residues on the N-terminal side of tyrosine. The presence of Ala replacing PKA E170 in all the members of the Src family may reflect their reduced preference for acidic residues compared to Syk and receptor PTKs, which have an Arg in that position; it is not consistent however with the presence of arginine in CSK and Abl as well, both even less acidophilic than Src themselves. More reliable information may be drawn from the nature of the residue homologous to PKA L205: this is invariably Met in receptor kinases, which fits in with their preference for a hydrophobic residue at $n + 1$, while it is Thr in most non-receptor PTKs; since threonine is likely to be hydrated this could account for the observation that non-receptor PTKs often select small or hydrophilic residues at $n + 1$. The structural basis for the selection of a hydrophobic residue at position $n - 1$ by Src, Abl and most non-receptor PTKs (with the exception of Syk) remains unknown.

10. Open questions and perspectives

While a considerable amount of information is available about the consensus sequences of many Ser/Thr protein kinases and the molecular features underlying the sequence specificity of these enzymes, little is known about the local determinants and the actual relevance of sequence specificity in the case of protein tyrosine kinases. Our knowledge of the mode of binding of phosphoacceptor substrates into the catalytic site of protein kinases is still based on indirect inference from the crystal structure of PKA complexed with inhibitory pseudosubstrate peptide, while no protein kinase crystals with a bona fide bound phosphorylatable substrate have been described to date. Obtaining and solving the structure of such protein kinase crystals with bound genuine peptide substrates (rather than inhibitors or substrate analogs) will provide an invaluable breakthrough in the detailed understanding of the mode of binding of phosphoacceptor targets by protein kinases. This would be of critical importance in the case of protein tyrosine kinases, which do not appear to use the same substrate recognition elements as Ser/Thr protein kinases. Additional clues for mapping the spatial and chemical features that dictate specificity can be provided by new mutational studies aimed at altering either sequence or active site specificities of both Ser/Thr and Tyr specific protein kinases. A comprehensive knowledge of these features will allow to predict the site specificity of any kinase from its primary structure assuming that the tertiary structure conforms to the canonical one.

It can be expected in this respect that the crystal structure of more and more protein kinases will become available in the near future: this will help to refine our knowledge about the mode of substrate binding, especially for those kinases whose specificity determinants are poorly understood and/or whose activity appears to rely on the integrity of the native protein substrate(s). It may be not unrealistic in this respect to look forward to the crystallization of protein kinases complexed with their substrates through modules, like the SH2, SH3 and PTB domains, that are outside their catalytic core.

One can expect that recently developed methodologies exploiting peptide libraries will become more widely used and gradually replace the traditional approaches based on the synthesis of series of individual peptides derived from natural phosphoacceptor sites. The potential of the peptide libraries is enormous since in principle they can provide a huge amount of information in a single experiment, including optimal sequence and nature and position of both positive and negative determinants. As discussed above (Section 5.1), however, there are practical problems and theoretical limitations of peptide libraries that may hamper the collection of reliable data and make their interpretation difficult. It is to be hoped that efforts to optimize peptide libraries will be aimed at increasing their sensitivity and rendering

simpler and more straightforward the assessment of real, rather than 'virtual' sequences that are efficiently phosphorylated. Sensitivity can be enormously increased by reducing or abolishing degeneracy at position(s) where the absolute requirement of definite residue(s) has been already established. This has been done e.g. with oriented libraries for Pro-directed kinases in which the motif S-P was maintained constant [11] and can be done with other kinases exhibiting a clear-cut consensus if one wishes to gain additional information about the residues positively or negatively selected at other positions.

While the completeness of information provided by peptide libraries is decreased by reducing randomization, this shortcoming is amply compensated by increased sensitivity and other technical advantages. Moreover, lack of thoroughness can be obviated by adopting iterative approaches in which partially degenerate sublibraries are successively analyzed. Thus, using sublibraries with few (1-3) defined positions where known aminoacids are systematically changed, and/or reducing the number of aminoacids tested at the defined positions (e.g. by using only 6 to 10 representative residues instead of all aminoacids) the total number of peptides becomes manageable and in principle they can be kept separate during synthesis and individual assays with the additional advantage of not having to analyze the sequence of the phosphorylated derivatives (or a mixture of them) after the phosphorylation reaction. Sublibraries of peptides synthesized as individual spots on a cellulose paper support or on separate discs of hydrophylic matrix have been recently constructed and successfully used for delineating subtle differences between PKA and PKG consensus sequences [15] and for determining the specificity of TGFβ receptor protein kinase [14] respectively. The iterative approach using cellulose paper libraries [15] looks especially practical and promising: these were composed of arrays of 400 octapeptide pools that contained two defined positions while the other six position were either occupied by mixtures of all 20 natural aminoacids or by fixed aminoacids selected on the basis of the results with more degenerate sublibraries. Immersion of the cellulose paper in the protein kinase assay medium generated radiolabeled spots where phosphoacceptor peptide(s) were present, and the intensity of these spots reflected the phosphorylation efficiency by the protein kinase tested [15]. In fact, these maps could be exploited as 'fingerprints' of kinase specificity.

So much effort to unravel the consensus sequences and the specificity determinants of the members of a single family of enzymes may appear excessive. It should be remembered, however, that protein kinases are encoded by about 6% of the whole yeast genome and have the task of phosphorylating and regulating perhaps half of all the proteins present in a eukaryotic cell, thus influencing nearly every cellular function. Therefore detailed knowledge of the features by which phosphoacceptor sites are specifically recognized by diverse protein kinases will shed light on fundamental aspects of signal transduction and cell regulation.

11. Note added in proof

Specificity analysis based on oriented peptide library has been recently extended to additional Ser/Thr protein kinases [205] providing the following virtual optimal sequences: phosphorylase kinase [FRMM*S*FFLF]; CaM kinase II [KRQQ*S*FDLF]; CK1δ [FDTG*S*IIIF]; CK1γ [YDAA*S*IIIF]; CK2 [EDEE*S*EDEE]; ERK1 [TGPL*S*PGPF]; Cdk5 [KHHK*S*PRHR].

Acknowledgements

We thank the many investigators who provided us with reprints and preprints of their work, and David P. Leader for linguistic advice. Work in the authors' laboratory is supported by grants from AIRC, Italian Ministero della Sanità (Progetto AIDS) and CNR (target project ACRO and grant 95.02920.CT14).

References

[1] Kemp, B.E. and Pearson, R.B. (1990) Trends Biochem. Sci. 15, 342–346.
[1a] Pearson, R.B. and Kemp, B.E. (1991) Meth. Enzymol. 200, 62–81.
[2] Kennelly, P.J. and Krebs, E.G. (1991) J. Biol. Chem. 266, 15555–15558.
[3] Schlessinger, J. (1994) Curr. Opin. Genet. Dev. 4, 25–30.
[4] Haslam, R.J, Kolde, H.B. and Hemmings, B.A. (1993) Nature 363, 309–310.

[5] Mayer, B.J., Ren, R., Clark, K.L. and Baltimore, D. (1993) Cell 71, 359–362.
[6] Van der Geer, P. and Pawson, T. (1995) Trends Biochem. Sci. 20, 277–280.
[7] Eck, M.J. (1995) Structure 3, 421–424.
[8] Carmel, G. and Kuret, J. (1992) Anal. Biochem. 203, 274–280.
[9] Till, J.H., Annan, R.S., Carr, S.A. and Miller, W.T. (1994) J. Biol. Chem 269, 7423–7428.
[10] Wu, J., Ma, Q.N. and Lam, K.S. (1994) Biochemistry 33, 14825–14833.
[11] Songyang, Z., Blechner, S., Hoagland, N., Hoekstra, M.F., Piwnica-Worms, H. and Cantley, L.C. (1994) Curr. Biol. 4, 973–982.
[12] Songyang, Z., Carraway III, K.L., Eck, M.J., Harrison, S.C., Feldman, R.A., Mohammadi, M., Schlessinger, J., Hubbard, S.R., Smith, D.P., Eng, C., Lorenzo, M.J., Ponder, B.A.J., Mayer, B.J. and Cantley, L.C. (1995) Nature 373, 536–539.
[13] Lam, K.S., Wu, J. and Lou, Q. (1995) Int. J. Peptide Protein Res. 45, 587–592.
[14] Luo, K., Zhou, P. and Lodish, H.F. (1995) Proc. Natl. Acad. Sci. USA 92, 11761–11765.
[15] Tegge, W., Frank, R., Hofmann, F. and Dostmann, W.R.G. (1995) Biochemistry 34, 10569–10577.
[16] Hunter, T. (1994) Seminars Cell. Biol. 5, 367–376.
[17] Hubbard, M.J. and Cohen, P. (1993) Trends Biochem. Sci. 18, 172–177.
[18] Rubin, C.S. (1994) Biochim. Biophys. Acta 1224, 467–479.
[19] Kwon, Y-G, Mendelow, M., Srinivasan, J., Lee, T.R., Pluskey, S., Salerno, A. and Lawrence, D.S. (1993) J. Biol. Chem. 268, 10713–10716.
[20] Meggio, F., Donella-Deana, A., Ruzzene, M., Brunati, A.M., Cesaro, L., Guerra, B., Meyer, T., Mett, H., Fabbro, D., Furet, P., Dobrowolska, G. and Pinna, L.A. (1995) Eur. J. Biochem. 234, 317–322.
[21] Kwon, Y-G., Mendelow, M. and Lawrence, D.S. (1994) J. Biol. Chem. 269, 4839–4844.
[22] Lee, T.R., Niu, J. and Lawrence, D.S. (1995) J. Biol. Chem. 270, 5375–5380.
[23] Lee, T.R., Till, J.H., Lawrence, D.S. and Miller, W.T (1995) J. Biol. Chem. 270, 27022–27026.
[24] Wood, J.S., Yan, X., Mendelow, M., Corbin, J.D., Francis, S.H. and Lawrence, D.S. (1996) J. Biol. Chem. 271, 174–179.
[25] Hanks, S.K. and Hunter, T. (1995) FASEB J. 9, 576–596.
[26] Alex, L.A. and Simon, M.J. (1994) Trends Genet. 10, 133–136.
[27] Hanks, S.K., Quinn, A.M. and Hunter, T. (1988) Science 241, 42–52.
[28] Hubbard, S.R., Wei, L., Ellis, L. and Hendrickson, W.A. (1994) Nature 372, 746–754.
[29] Taylor, S.S., Radio-Andzelm, E. and Hunter, T. (1995) FASEB. J. 9, 1255–1266.
[30] Meggio, F., Boulton, A.P., Marchiori, F., Borin, G., Lennon, D.P.W., Calderan, A. and Pinna, L.A. (1988) Eur. J. Biochem. 177, 281–284.
[31] Leighton, I.A., Dalby, K.N., Caudwell, F.B., Cohen, P.T.W. and Cohen, P. (1995) FEBS Lett. 375, 289–293.
[32] Dale, S., Wilson, W.A., Edelman, A.M. and Hardie, D.G. (1995) FEBS Lett. 361, 191–195.
[33] Kamijo, M., Yasuda, H., Yau, P.M., Yamashita, M., Nagahama, Y. and Ohba, Y. (1992) Peptide Res. 5, 281-285
[33a] Lindberg, R.A., Fischer, W.H. and Hunter, T. (1993) Oncogene 8, 351–359
[33b] Wakim, B.T. and Aswad, G.D. (1994) J. Biol. Chem. 269, 2722–2727.
[34] Huang, J.M., Wei, Y.F., Kim, Y.H., Osterberg, L. and Matthews, H.R. (1991) J. Biol. Chem. 266, 9023–9031.
[35] Crovello, C.S., Furie, B.C. and Furie, B. (1995) Cell 82, 279–286.
[36] Lindberg, R.A., Quinn, A.M. and Hunter, T. (1992) Trends Biochem. Sci. 17, 114–119.
[37] Letwin, K., Mizzen, L., Motro, B., Ben-David, Y., Bernstein, A. and Pawson, T. (1992) EMBO J. 11, 3521–3531.
[38] Lauzè, E., Stoelcker, B., Luca, F.C., Weiss, E., Schutz, A.R. and Winey, M. (1995) EMBO J. 14, 1655–1663.
[39] Douville, E.M., Afar, D.E., Howell, B.W., Letwin, K., Tannock, L., Ben-David, Y., Pawson, T. and Bell, J.C. (1992) Mol. Cell. Biol. 12, 2681–2689.
[40] Hoekstra, M.F., Dhillon, N., Carmel, G., DeMaggio, A.J., Lindberg, R.A., Hunter, T. and Kuret, J. (1994) Mol. Biol. Cell 5, 877–886.
[41] Mueller, P.R., Coleman, T.R., Kumagai, A. and Dunphy, W.G. (1995) Science 270, 86–89.
[42] Honda, R., Ohba, Y. and Yasuda, H. (1992) Biochem. Biophys. Res. Commun. 186, 1333–1338.
[43] Huang, C-Y., Yuan, C-J., Luo, S. and Graves, D.J. (1994) Biochemistry 33, 5877-5883.
[43a] Stern, D.F., Zheng, P., Beidler, D.R. and Zerillo, C. (1991) Mol. Cell. Biol. 11, 987–1001.
[44] L'Allemain, G., Her, J., Wu, J., Sturgill, T.W. and Weber, M.J. (1992) Mol. Cell. Biol. 5, 2222–2229.
[45] Nakielny, S., Cohen, P., Wu, J. and Sturgill, T. (1992) EMBO J. 11, 2123–2129.
[46] Rossomando, S., Wu, J., Weber, M.J. and Sturgill, T.W. (1992) Proc. Natl. Acad. Sci. USA 89, 5221–5225.
[47] Her, J-H., Lakhani, S., Zu, K., Vila, J., Dent, P., Sturgill, T.W. and Weber, M.J. (1993) Biochem J. 296, 25–31.
[48] Songyang, Z. and Cantley, L.C. (1995) Trends Biochem. Sci. 20, 470–475.
[49] Donella-Deana, A., James, P., Staudenmann, W., Cesaro, L., Marin, O., Brunati, A.M., Ruzzene, M. and Pinna, L.A. (1996) Eur. J. Biochem. 235, 18–25.
[50] Szallasi, Z., Denning, M.F., Chang, E-Y., Rivera, J., Yuspa, S.H., Lehel, C., Olah, Z., Anderson, W.B. and Blumberg, P.M. (1995) Biochem. Biophys. Res. Commun. 214, 888–894.
[51] Carrera, A.C., Paradis, H., Borlado, L.R., Roberts, T.M. and Martinez, C. (1995) J. Biol. Chem. 270, 3385–3391.

[52] Kavanaugh, W.M., and Williams, L.T. (1994) Science 266, 1862–1865.
[53] Bork, P. and Margolis, B. (1995) Cell 80, 93–104.
[54] Songyang, Z., Margolis, B., Chaudhuri, M., Shoelson, S.E. and Cantley, L.C. (1995) J. Biol. Chem. 270, 14863–14866.
[55] Voliovich, H., Schindler, D.G., Hadari, Y.R., Taylor, S.I., Accili, D. and Zick, Y. (1995) J. Biol. Chem. 270, 18083–18087.
[56] Alvarez, C.V., Shon, K-J., Miloso, M. and Beguinot, L. (1995) J. Biol. Chem. 270, 16271–16276.
[57] Ikebe, M., Reardon, S., Schwonek, J.P., Sanders, C.R.II and Ikebe, R. (1994) J. Biol. Chem. 269, 28165–28172.
[58] Krueger, J.K., Padre, R.C. and Stull, J.T. (1995) J. Biol. Chem. 270, 16848-16853.
[58a] Srinivasan, J., Koszelak, M., Mendelow, M., Kwon, Y-G. and Lawrence, D.S. (1995) Biochem. J. 309, 927–931.
[59] Perich, J.W., Meggio, F., Reynolds, E.C., Marin, O. and Pinna, L.A. (1992) Biochemistry 31, 5893–5897.
[60] Brunati, A.M., Donella-Deana, A., Ruzzene, M., Marin, O. and Pinna, L.A. (1995) FEBS Lett. 367, 149–152.
[61] Zetterqvist, O., Ragnarsson, U., Humble, H., Berglund, L. and Engstrom, L. (1976) Biochem. Biophys. Res. Commun. 70, 696–703.
[62] Kemp, B.E., Graves, D.J., Benjamini, E. and Krebs, E.G. (1977) J. Biol. Chem. 252, 4888–4894.
[63] Feramisco, J.R., Glass, D.B. and Krebs, E.G. (1980) J. Biol. Chem. 255, 4240–4245.
[64] Meggio, F., Chessa, G., Borin, G., Pinna, L.A. and Marchiori, F. (1981) Biochim. Biophys. Acta 662, 94–101.
[65] Leader, D.P., Donella-Deana, A., Marchiori, F., Purves, F.C. and Pinna, L.A. (1991) Biochim. Biophys. Acta 1091, 426–431.
[66] Mitchell, R.D., Glass, D.B., Wong, C-W, Angelos, K.L. and Walsh, D.A. (1995) Biochemistry 34, 528–534.
[67] Knighton, D.R., Zheng, J., Ten Eyck, L.F., Xuong, N-H., Taylor, S.S. and Sowadski, J.M. (1991) Science 253, 414–420.
[68] Glass, D.B., El-Maghrabi, M.R. and Pilkis, S.J. (1986) J. Biol. Chem. 261, 2987–2993.
[69] Chessa, G., Borin, G., Marchiori, F., Meggio, F., Brunati, A.M. and Pinna, L.A. (1983) Eur. J. Biochem. 135, 609–614.
[70] Chang, X-B., Tabcharani, J.A., Hou, Y-X., Jensen T.J., Kartner, N., Alon, N., Hanrahan, J.W. and Riordan, J.R. (1993) J. Biol. Chem. 268, 11304–11311.
[71] Glass, D.B. (1990) in Peptides and Protein Phosphorylation (Kemp, B.E. ed.) pp 209–238, CRC Press Inc., Boca Raton, FL.
[72] Yan, X., Corbin, J.D., Francis, S.H. and Lawrence, D.S. (1996) J. Biol. Chem. 271, 1845–1848.
[73] Ferrari, S., Marchiori, F., Borin, G. and Pinna, L.A. (1985) FEBS Lett. 184, 72–77.
[74] Ferrari, S., Marchiori, F., Marin, O. and Pinna, L.A. (1987) Eur. J. Biochem. 163, 481–487.
[74a] Woodgett, J.R., Gould, K.L. and Hunter, T. (1986) Eur. J. Biochem. 161, 177–184.
[75] Graff, J.M., Stumpo, D.J. and Blackshear, P.J. (1989) J. Biol. Chem. 264, 11912–11919.
[76] Kielbassa, K., Müller, H-J., Meyer, H.E., Marks, F. and Gschwendt, M. (1995) J. Biol. Chem. 270, 6156–6162.
[77] Ferrari, S. and Pinna, L.A (1987) Biochem. Biophys. Res. Commun. 144, 1324–1331.
[78] Palmer, R.H. and Parker, P.J. (1995) Biochem. J. 309, 315–320.
[79] Palmer, R.H., Schönwaßer, D.C., Rahman, D., Pappin, D.J.C., Herget, T. and Parker, P.J. (1996) FEBS Lett. 378, 281–285.
[80] Ingley, E. and Hemmings, B.A. (1995) in The Protein Kinase Facts Book I (Hardie, G. and Hanks, S., eds.), pp. 95–98, Academic Press, London.
[81] Cross, D.A.E., Alessi, D.R., Cohen, P., Andejelkovich, M. and Hemmings, B.A. (1995) Nature 378, 785–789.
[82] Flotow, H. and Thomas, G. (1992) J. Biol. Chem. 267, 3074–3078.
[83] Donella-Deana, A., Lavoinne, A., Marin, O., Pinna, L.A. and Cohen, P. (1993) Biochim. Biophys. Acta 1178, 189–193.
[84] Stokoe, D., Caudwell, B., Cohen, P.T.W. and Cohen, P. (1993) Biochem. J. 296, 843–849.
[85] Hanson, P.I and Schulman, H. (1992) Annu. Rev. Biochem. 61, 559–601.
[86] Zhi, G., Herring, B.P. and Stull, J.T. (1994) J. Biol. Chem. 269, 24723–24727.
[87] Huang, C-Y., Yuan, C-J., Blumenthal, D.K. and Graves, D.J. (1995) J. Biol. Chem. 270, 7183–7188.
[88] Sarno, S., Vaglio, P., Issinger, O-G. and Pinna, L.A. (1996) J. Biol. Chem. 271, 10595–10601.
[89] Chan, K-F. J., Hurst, M.O. and Graves, D.J. (1982) J. Biol. Chem. 257, 3655–3659.
[90] Lee, J.C., Kwon, Y-G., Lawrence, D.S. and Edelman, A.M. (1994) Proc. Natl. Acad. Sci. USA 91, 6413–6417.
[91] Proud, C.G., Colthrust, D.R., Ferrari, S. and Pinna, L.A. (1991) Eur. J. Biochem. 195, 771–779.
[92] Purves, F.C., Donella-Deana, A., Marchiori, F., Leader, D.P. and Pinna, L.A. (1986) Biochim. Biophys. Acta 889, 208–215.
[92a] Peeper, D.S., Parker, L.L., Ewen, M.E., Toebes, M., Hall, F.L., Xu, M., Zantema, A., van der Eb, A.J. and Piwnica-Worms, H. (1993) EMBO J. 12, 1947–1954.
[93] Mathias, S., Younes, A., Kan, C., Orlow, I., Joseph, C. and Kolesnick, R.N. (1993) Science 259, 519–522.
[94] Joseph, C.K., Buyn, H-S., Bittman, R. and Kolesnick, R.N. (1993) J. Biol. Chem. 268, 20002–20006.
[95] Pinna, L.A. (1990) Biochim. Biophys. Acta 1054, 267–284.
[96] Pinna, L.A. (1994) Cell. Mol. Biol. Res. 40, 383–390.
[97] Meggio, F., Marchiori, F., Borin, G., Chessa, G. and Pinna, L.A. (1984) J. Biol. Chem. 259, 14576–14579.
[98] Marin, O., Meggio, F., Marchiori, F., Borin, G. and Pinna, L.A. (1986) Eur. J. Biochem. 160, 239–244.
[99] Kuenzel, E.A., Mulligan, J.A., Sommercorn, J. and Krebs, E.G. (1987) J. Biol. Chem. 262, 9136–9140.
[100] Marchiori, F., Meggio, F., Marin, O., Borin, G., Calderan,

A., Ruzza, P. and Pinna, L.A. (1988) Biochim. Biophys. Acta 971, 332–338.
[101] Meggio, F., Perich, J.W., Meyer, H.E., Hoffmann-Posorske, E., Lennon, D.P.W., Johns, R.B. and Pinna, L.A. (1989) Eur. J. Biochem. 186, 459–464.
[102] Pinna, L.A., Meggio, F. and Marchiori, F. (1990) in Peptides and Protein Phosphorylation (Kemp, B.E., ed.), pp. 145–169, CRC Press Inc., Boca Raton, FL.
[103] Meggio, F., Marin, O. and Pinna, L.A. (1994) Cell. Mol. Biol. Res. 40, 401–409.
[104] Lasa-Benito, M., Marin, O., Meggio, F. and Pinna, L.A. (1996) FEBS Lett. 382, 149–152.
[105] Meggio, F., Perich, J.W., Reynolds, E.C. and Pinna, L.A. (1991) FEBS Lett. 279, 307–309.
[106] Marin, O., Meggio, F., Daretta, G. and Pinna, L.A. (1992) FEBS Lett. 301, 111–114.
[107] Hu, B., Coulson, L., Moyer, B. and Price, P.A. (1995) J. Biol. Chem. 270, 431–436.
[108] Mercier, J.C. (1981) Biochemie 63, 1–17.
[109] Zhai, L., Graves, P.R., Robinson, L.C., Italiano, M., Culbertson, M.R., Rowles, J., Cobb, M.H., DePaoli-Roach, A.A. and Roach, P.J. (1995) J. Biol. Chem. 270, 12717–12724.
[110] Fish, K.J., Cegielska, A., Getman, M.E., Landes, G.M. and Virshup, D.M. (1995) J. Biol. Chem. 270, 14875–14883.
[111] Meggio, F., Donella-Deana, A. and Pinna, L.A. (1979) FEBS Lett. 106, 76-80.
[111a]Flotow, H., Graves, P.R., Wang, A, Fiol, C.J., Roeske, R.W. and Roach, P.J. (1990) J. Biol. Chem. 265, 14264–14269.
[112] Meggio, F., Perich, J.W., Reynolds, E.C. and Pinna, L.A. (1991) FEBS Lett. 283, 303–306.
[113] Meggio, F., Perich, J.W., Marin, O. and Pinna, L.A. (1992) Biochem. Biophys, Res. Commun. 182, 1460–1465.
[114] Agostinis, P., Marin, O., James, P., Hendrix, P., Merlevede, W., Vandenheede, J.R. and Pinna, L.A. (1992) FEBS Lett. 305, 121–124.
[115] Desdouits, F., Cohen, D., Nairn, A.C., Greengard, P. and Girault, J-A. (1995) J. Biol. Chem. 270, 8772–8778.
[116] Marin, O., Meggio, F., Sarno, S., Andretta, M. and Pinna, L.A. (1994) Eur. J. Biochem. 223, 647–653.
[117] Graves, P.R. and Roach, P.J. (1995) J. Biol. Chem. 270, 21689–21694.
[118] Picton, C., Woodgett, J., Hemmings, B. A. and Cohen, P. (1982) FEBS Lett. 150, 191–196.
[119] DePaoli-Roach, A.A., Ahmad, Z., Camici, M., Lawrence, J.C. Jr. and Roach, P.J. (1983) J. Biol. Chem. 258, 10702–10709.
[120] Zhang, W., DePaoli-Roach, A.A. and Roach, P.J. (1993) Arch. Biochem. Biophys. 304, 219–225.
[121] Roach, P.J. (1991) J. Biol. Chem. 266, 14139–14142.
[122] Fiol, CJ., Mahrenholz, A.M., Wang, Y., Roeske, R.W. and Roach, P.J. (1987) J. Biol. Chem. 262, 14042–14048.
[123] Skurat, A.V. and Roach, P.J. (1995) J. Biol. Chem. 270, 12491–12497.
[124] Premont, R.T., Inglese, J. and Lefkowitz, R.J. (1995) FASEB J. 9, 175–182.
[125] Palczewski, K., McDowell, J.H. and Hargrave, P.A. (1988) Biochemistry 27, 2306–2313.
[126] Kelleher, D.J. and Johnson, G.L. (1990) J. Biol. Chem. 265, 2632–2639.
[127] Benovic, J.L., Onorato, J.J., Lohse, M.J., Dohlman, H.G., Staniszewski, C., Caron, M.G. and Lefkowitz, R.J. (1990) Br. J. Clin. Pharmacol. 30, 3S–12S.
[128] Onorato, J.J., Palczewski, K., Regan, J.W., Caron, M.G., Lefkowitz, R.J. and Benovic, J.L. (1991) Biochemistry 30, 5118–5125.
[129] Eason, M.G., Moreira, S.P. and Liggett, S.B. (1995) J. Biol. Chem. 270, 4681–4688.
[130] Pearson, R.B., Dennis, P.B., Han, J.W., Williamson, N.A., Kozma, S.C., Wettenhall, R.E. and Thomas, G. (1995) EMBO J. 14, 5279–5287.
[131] Mullinax, T.R., Stepp, L.R., Brown, J.R. and Reed, L.J. (1985) Arch. Biochem. Biophys. 243, 655–659.
[132] Paxton, R., Kuntz, M. and Harris, R.A. (1986) Arch. Biochem. Biophys. 244, 187–201.
[133] Hunter, T. (1995) Cell 83, 1–4.
[134] Hartley, K.O., Gell, D., Smith, G.C.M., Zhang, H., Divecha, N., Connelly, M.A., Admon, A., Lees-Miller, S.P., Anderson, C.W. and Jackson, S.P. (1995) Cell 82, 849–856.
[135] Lees-Miller, S.P. and Anderson, C.W. (1991) Cancer Cells 9, 341–346.
[136] Watanabe, F., Teraoka, H., Iijima, S., Mimori, T. and Tsukada, K. (1994) Biochim. Biophys. Acta 1223, 255–260.
[137] Geahlen, R.L. and Harrison, M.L. (1990) in Peptides and Protein Phosphorylation (Kemp, B.E., ed.), pp. 239–253, CRC Press, Boca Raton, FL.
[138] Marin, O., Donella-Deana, A., Brunati, A.M., Fischer, S. and Pinna, L.A. (1991) J. Biol. Chem. 266, 17798–17803.
[139] Donella-Deana, A., Marin, O., Brunati, A.M., Cesaro, L., Piutti, C. and Pinna, L.A. (1993) FEBS Lett. 330, 141–145.
[140] Budde, R.J.A., Obeyesekere, N.U., Ke, S. and McMurray, J.S. (1995) Biochim. Biophys. Acta 1248, 50–56.
[141] Casnellie, J.E. and Krebs, E.G. (1984) Adv. Enzyme Regul. 22, 501–515.
[142] Donella-Deana, A., Brunati, A.M., Marchiori, F., Borin, G., Marin, O. and Pinna, L.A. (1990) Eur. J. Biochem. 194, 773–777.
[143] Donella-Deana, A., Varro, A., Dockray, G.J. and Pinna, L.A. (1991) Biochim. Biophys. Acta 1095, 75–77.
[144] Donella-Deana, A., Stone, S.R. and Pinna, L.A. (1991) Eur. J. Biochem. 201, 501–505.
[145] Ruzzene, M., Donella-Deana, A., Marin, O., Perich, J.W., Ruzza, P., Borin, G., Calderan, A. and Pinna, L.A. (1993) Eur. J. Biochem. 211, 289–295.
[146] Okada, M., Shigeyuki, N., Yamanashi, Y., Yamamoto, T. and Nakagawa, H. (1991) J. Biol. Chem. 266, 24249–24252.
[147] Levine, B.A., Clack, B. and Ellis, L. (1991) J. Biol. Chem. 266, 3565–3570.
[148] Keane, N.E., Chavanieu, A., Quirk, P.G., Evans, J.S., Levine, B.A., Calas, B., Wei, L. and Ellis, L. (1994) Eur. J. Biochem 226, 525–536.

[149] Shoelson, S.E., Chatterjee, S., Chandhuri, M. and White, M.F. (1992) Proc. Natl. Acad. Sci. USA 89, 2027–2031.
[150] Xu, B., Bird, V.G. and Miller, W.T. (1995) J. Biol. Chem. 270, 29825–29830.
[151] Price, N.T., Redpath, N.T., Severinov, K.V., Campbell, D.G., Russell, J.M. and Proud, C.G. (1991) FEBS Lett. 282, 253–258.
[152] Sahyoun, N., McDonald, O.B., Farrell, F. and Lapetina, E.G. (1991) Proc. Natl. Acad. Sci. USA 88, 2643–2647.
[153] Miyano, O., Kameshita, I. and Fujisawa, H. (1992) J. Biol. Chem. 267, 1198–1992.
[154] Weekes, J., Ball, K.L., Caudwell, F.B. and Hardie, D.G. (1993) FEBS Lett. 334, 335–339.
[155] Yang, S-D., Huang, T-J. and Soderling, T.R. (1994) J. Biol. Chem. 269, 29855–29859.
[156] Morrice, N.A., Gabrielli, B., Kemp, B.E. and Wettenhall, R.E.H. (1994) J. Biol. Chem. 269, 20040–20046.
[157] Lu, K.P., Kemp, B.E. and Means, A.R. (1994) J. Biol. Chem. 269, 6603–6607.
[158] Fry, A.M., Schultz, S.J., Bartek, J. and Nigg, E.A. (1995) J. Biol. Chem. 270, 12899–12905.
[159] Gonzales, F.A., Raden, D.L. and Davis, R.J. (1991) J. Biol. Chem. 266, 22159–22163.
[160] Beaudette, K.N., Lew, J. and Wang, J.H. (1993) J. Biol. Chem. 268, 20825–20830.
[161] Kidd, V.J. (1995) in The Protein Kinase Facts Book I (Hardie, G. and Hanks, S. eds.) pp. 260–265, Academic Press Limited, London, S. Diego.
[162] Pullen, N. and Akhtar, M. (1994) Biochemistry 33, 14536–14542.
[163] Garcia, P., Shoelson, S.E., George, S.T., Hinds, D.A., Goldberg, A.R. and Miller, W.T. (1993) J. Biol. Chem. 268, 25146–25151.
[164] Bellis, S.L., Miller, J.T. and Turner, C.E. (1995) J. Biol. Chem 270, 17437–17441.
[165] Mayer, B.J., Hirai, H. and Sakai, R. (1995) Curr. Biol. 5, 296–305.
[166] Duyster, J., Baskaran, R. and Wang, J.Y. (1995) Proc. Natl. Acad. Sci. USA 92, 1555–1559.
[167] Brunati, A.M., Ruzzene, M., James, P., Guerra, B. and Pinna, L.A. (1995) Eur. J. Biochem. 229, 164–170.
[168] Ruzzene, M., Brunati, Marin, O., A.M., Donella-Deana, A., and Pinna, L.A. (1996) Biochemistry 35, 5327–5332.
[169] Pinna, L.A. and Donella-Deana, A. (1994) Biochim. Biophys. Acta 1222, 415–431.
[170] Donella-Deana, A., Krinks, M.H., Ruzzene, M., Klee, C. and Pinna, L.A. (1994) Eur. J. Biochem. 219–109–117.
[171] Agostinis, P., Goris, J., Waelkens, E., Pinna, L.A., Marchiori, F. and Merlevede, W. (1987) J. Biol. Chem. 262, 1060–1064.
[172] Agostinis, P., Goris, J., Pinna, L.A., Marchiori, F., Perich, J.W., Meyer, H.E. and Merlevede, W. (1990) Eur. J. Biochem. 189, 235–241.
[173] Donella-Deana, A., McGowan, C.H., Cohen, P., Marchiori, F., Meyer, H.E. and Merlevede, W. (1990) Biochim. Biophys. Acta 1051, 199–202.
[174] Ben-Levy, R., Leighton, I.A., Doza, Y.N., Attwood, P., Morrice, N., Marshall, C.J. and Cohen, P. (1995) EMBO J. 14, 5920–5930.
[175] Pinna, L.A., Agostinis, P., Donella-Deana, A. and Marchiori, F. (1989) Adv. Protein Phosphatases 5, 51–74.
[176] Körner, C., Herzog, A., Weber, B., Rosorius, O., Hemer, F., Schmidt, B. and Braulke, T., (1994) J. Biol. Chem. 269, 16529–16532.
[177] Ulloa, L., Dombradi, V., Diaz-Nido, J., Szucs, K., Gergely, P., Friedrich, P. and Avila, J. (1993) FEBS Lett. 330, 85–89.
[178] Small, D., Chou, P.J. and Fasman, G.D. (1977) Biochim. Biophys. Res. Commun. 79, 341–346.
[179] Pinna, L.A., Donella-Deana, A. and Meggio, F. (1979) Biochim. Biophys. Res. Commun. 87, 114–120.
[180] Tinker, D.A., Krebs, E.A., Feltham, I.C., Attah-Poku, S.K. and Ananthanarayanan, V.S. (1988) J. Biol. Chem. 263, 5024–5026.
[181] Ruzza, P., Calderan, A., Filippi, B., Biondi, B., Donella-Deana, A., Cesaro, L., Pinna, L.A. and Borin, G. (1995) Int. J. Peptide Protein Res. 45, 529–539.
[182] Park, I-K., Roach, P.J., Bondor, J., Fox, S.P. and DePaoli-Roach, A.A. (1994) J. Biol. Chem. 269, 944–954.
[183] Boldyreff, B., Meggio, F., Pinna, L.A. and Issinger, O-G. (1994) J. Biol. Chem. 269, 4827–4831.
[184] Dumas, C., Lascu, I., Morèra, S., Galser, P., Fourme, R., Wallet, V., Lacombe, M-L., Nèron, M. and Janin, J. (1992) EMBO J. 11, 3203–3208.
[185] Engel, M., Issinger, O-G., Lascu, I., Seib, T., Dooley, S., Zang, K.D. and Welter, C. (1994) Biochim. Biophys. Res. Commun. 199, 1041–1048.
[186] Knighton, D.R., Zheng, J., Ten Eyck, L.F., Ashford, V.A., Xuong, N-H., Taylor, S.S. and Sowadski, J.M. (1991) Science 253, 407–413.
[187] Bossemeyer, D., Engh, R.A., Kinzel, V., Ponstingl, H. and Huber, R. (1993) EMBO J., 12, 849–859.
[188] De Bondt, H.L., Rosenblatt, J., Jancarik, J., Jones, H.D., Morgan, D.O. and Kim, S-H. (1993) Nature 363, 595–602.
[189] Zhang, F., Strand, A., Robbins, D., Cobb, M.H. and Goldsmith, E.J. (1994) Nature 367, 704–711.
[190] Hu, S-H., Parker, M.W., Lei, J.Y., Wilce, M.C.J., Benian, G.M. and Kemp, B.E. (1994) Nature 369, 581–584.
[191] Xu, R.M., Carmel, G., Sweet, R.M., Kuret, J. and Cheng, X. (1995) EMBO J. 14, 1015–1023.
[192] Buechler, J.A. and Taylor, S.S. (1990) Biochemistry 29, 1937–1943.
[193] Gibbs, C.S. and Zoller, M.J. (1991) Biochemistry 30, 5329–5334.
[194] Gibbs, C.S. and Zoller, M.J. (1991) J. Biol. Chem. 266, 8923–8931.
[195] Cox, S. and Taylor, S.S. (1995) Biochemistry 34, 16203–16209.
[196] Herring, B.P., Gallagher, P.J. and Stull, J.T. (1992) J. Biol. Chem. 267, 25945–25950.
[197] Gallagher, P.J., Herring, B.P., Trafny, A., Sowadski, J. and Stull, J.T. (1993) J. Biol. Chem. 268, 26578–26582.

[198] Dobrowolska, G., Meggio, F., Marin, O., Lozeman, F.J., Li, D., Pinna, L.A. and Krebs, E.G. (1994) FEBS Lett. 355, 237–241.
[199] Sarno, S., Boldyreff, B., Marin, O., Guerra, B., Meggio, F., Issinger, O-G. and Pinna, L.A. (1995) Biochem. Biophys. Res. Commun. 206, 171–179.
[200] Herberg, F.W. and Taylor, S.S. (1993) Biochemistry 32, 14015–14022.
[201] Wang, X. and Robinson, P.J. (1995) J. Neurochem. 65, 595–604.
[202] Longenecker, K.L., Roach, P.J. and Hurley, T.D. (1996) J. Mol. Biol. 257, 618–631.
[203] Graves, P.R., Haas, D.W., Hagedorn, C.H., DePaoli-Roach, A.A. and Roach, P.J. (1993) J. Biol. Chem. 268, 6394–6401.
[204] Palczewski, K., Ohguro, H., Premont, R.T. and Inglese, J. (1995) J. Biol. Chem. 270, 15294–15298.
[205] Songyang, Z., Lu, K.P., Kuan, Y., Tsai, L.-H., Filhol, O., Cochet, C., Soderling, T.R., Bartleson, C., Graves, D.J., Hoekstra, M.F., Blenis, J., Hunter, T. and Cantley, L.C. (1996) Mol. Cell. Biol., in press.