Identity of the putative serine-proteinase fold in proteins of the complement system with nine relevant crystal structures

Stephen J. PERKINS* and Kathryn F. SMITH

Department of Biochemistry and Chemistry, Royal Free Hospital School of Medicine, Rowland Hill Street, London NW3 2PF, U.K.

The serine-proteinase domain is responsible for the proteolytic events that occur during complement activation. The sequences of nine serine proteinases of known crystal structure were compared with the serine-proteinase sequences in the six complement proteins C1r, C1s, C2, factor B, factor I and factor D to assess the degree of structural homology of the latter with the crystal structures. All sequence insertions and deletions were

readily located at the protein surface. The internal location of disulphide bridges and the surface location of putative glycosylation sites are compatible with this structure. Secondary-structure predictions for the sequences were fully consistent with the crystal structures. It is concluded that the double subdomain β -sheet motif is retained in the complement sequences, but that localized differences are observed for factor I, C2 and factor B.

INTRODUCTION

Proteins in the complement system of immune defence are constructed from single or multiple copies of independently folded domains (Reid et al., 1986; Sim and Perkins, 1989). The serine-proteinase (SP) domain occurs once in each of C1r, C1s and C2, and in factors B, D and I. In the classical pathway of the complement cascade, the SP domains in the C1r dimer within the C1r₂C1s₂ tetramer become autoactivated during C1 activation by immune complexes. Each SP domain in C1r then activates that in C1s, which is then able to cleave C2 (and also C4). In the alternative pathway of activation, factor D has the analogous role of C1s, and cleaves and activates factor B. Cleaved C2 and factor B are able to cleave C3, and then C5, in a complex series of reactions to lead to the formation of the membrane-attack complex (Reid, 1986; Law and Reid, 1988). As a control mechanism, factor I is responsible for the proteolytic inactivation of C3b and C4b, using factor H and the C4b binding protein respectively as cofactors.

The general SP activities of these complement proteins have been established biochemically. However, unusual properties are displayed by all six proteins. The multidomain components C1r, C1s and factor I are activated by a single peptide cleavage at a conserved location at the N-terminus of the SP domain to yield a disulphide-linked two-chain structure in which the smaller chain is the SP domain. Factor I is unusual in that it is not normally inhibited by any known protein proteinase inhibitor, nor by the commonly used synthetic SP inhibitors. The SP domains of the five-domain proteins factor B and C2 are unusual in that the cleavage leading to proenzyme activation occurs at a site distant from the SP domain, which is then closely associated with the von-Willebrand-factor domain in a two-domain structure. Factors D and I are unusual in that they have no circulating inactive precursor forms. From the genetic standpoint it is interesting that the SP domain of C1r and C1s has an intron-less structure closely related to haptoglobin (Tosi et al., 1989), whereas the trypsin, chymotrypsin and elastase genes contain four to seven exons (Rogers, 1985), and the SP domain of factor B and probably C2 is coded for by eight exons (Campbell and Porter, 1983; Ishikawa et al., 1990).

The crystallographically known SP structure (as found in trypsin, α -chymotrypsin, elastase, kallikrein and other proteins) consists of two subdomains, each constructed from a six-stranded antiparallel β -sheet barrel structure (Chothia and Janin, 1982). The domains are linked by a short polypeptide hinge, and also by one disulphide bridge (chymotrypsinogen A) or two disulphide bridges (trypsinogen). The catalytic site is located between the two subdomains, and each subdomain is stabilized by further disulphide bridges within each one.

After the sequences of most of the complement SP domains were originally determined, it became recognized that sequence similarity is an insufficient condition to show that structural homology exists (Reeck et al., 1987; Lewin, 1987). The degree of structural identity between the SPs of known crystal structure and the six SP domains found in these proteins is required for structural studies on these complement proteins by solution scattering, Fourier-transform i.r. spectroscopy and other techniques. This identity has not yet been critically assessed, despite the unusual SP activities of these proteins and their diverse gene structures. Now that the factor D sequence is available to complete all six sequences (White et al., 1992), the comparison of relevant crystal structures with these SP sequences will establish the extent to which these SP domains can be modelled or analysed on the basis of these crystal structures. An initial sequence alignment has been reported by Greer (1990) for four full SP sequences and one partial one of the six in complement.

MATERIALS AND METHODS

Analysis of SP sequences

Secondary structure analyses of the SP domain were based on the following nine highly refined crystal structures: bovine α -chymotrypsin (Cohen et al., 1981; Wang et al., 1985; Tsukada and Blow, 1985), bovine β -trypsin (Bode and Schwager, 1975), rat trypsin (Sprang et al., 1987), *Streptomyces griseus* trypsin (Read and James, 1988), porcine elastase (Sawyer et al., 1978), human leukocyte elastase (An-Zhi et al., 1988; Navia et al., 1989), porcine kallikrein (Bode et al., 1983), rat tonin (Fujinaga and James, 1987) and rat mast-cell protease II (Remington et al., 1988). The Brookhaven crystal structure sequences were aligned

Abbreviation used: SP, serine proteinase.

* To whom correspondence and requests for reprints should be addressed.

	*	***	**	***	*****		***	*	*****	**	*
Factor D	RTHDGA	ITERLMCAE	SNRR	DSCKGSGGGLVCG		GVLEGVVTSSG	RVCG	NRKKPGIYTRVASYAAWIDSLVA			
Factor I	FYGNRFY	EKEMECAGTYDSI		DACKGSGGGLVQDANN		GLYGVGVSHGE	NKG	KPEFGVYTVKANYVDWISYHVGRRFISQYNV			
Factor B		DAQYAPGYDKVKDISEVTPRFLCTGGVSPYADPNTCRGDSGGLLVHKRSR				FIQVGVISHGVLVDCKNQKQKQVAPARDFHINLQVFLPHLKEKLQDE					
C2	VVSQEKTFMPLNLTDRREVVTQDFLCSG		TOEDESQCKGESGGAVFLERRFR			FFQVGLVSHGVLNPLNCLGSGADKNRSRKAPRSDKPPPRDFHINLFRMQPL					
C1r	HLRGKRN	MDVFSQNMFCAGHP	SLKQ	DSCKGSGGVFAVRDPN		TDRAWQVITVSWGI	GCS	RGYGFYTVKLVNVDWIKKEMEED			
C1s	VKVEKPT	ADAEAYVTFPMNICAGGE	KG M	DSCKGSGGGAFAVQDPNDKTKFYAAGLVSHGP			QCG	TYGLYTRKNYVDWIMTKQENSTPRED			
Cvs=Cvs/CHO		<u>C</u>	<u>8</u>	<u>9</u>	<u>7</u>	<u>C</u>	<u>Q</u>	<u>6</u>	<u>C</u>		

on the basis of topologically equivalent residues as defined in Read and James (1988) for chymotrypsin, bovine and *S. griseus* trypsin, tonin and kallikrein, in Remington et al. (1988) for α -chymotrypsin versus rat mast-cell protease II, in James et al. (1978) for α -chymotrypsin versus porcine elastase, in An-Zhi et al. (1988) for α -chymotrypsin versus human elastase, and by eye for rat trypsin versus bovine β -trypsin. This alignment is supported also by the multiparameter superimposition method of Sali and Blundell (1990) and an independent topological alignment by Greer (1990). The assignment of secondary-structure features is based on the DSSP program of Kabsch and Sander (1983) which was used to analyse objectively the structures of chymotrypsin, bovine and *S. griseus* trypsin, porcine elastase and tonin in terms of α -helices, 3_{10} helices, β -strands and turns (Fujinaga and James, 1987; Read and James, 1988).

Secondary-structure predictions

Consensus secondary-structure predictions are based on the Robson and Chou–Fasman methods (Garnier et al., 1978; Chou and Fasman, 1978). Both the original and updated parameter sets for the α -helix, β -sheet, β -turn and coil conformations were used in the Robson method (Garnier et al., 1978; Gibrat et al., 1987; Garnier and Robson, 1989), but only the 1978 set is used below. The nine SP predictions were averaged using FORTRAN software (PREDRB and PREDCF; Perkins et al., 1988). This produced the percentage of the sequences that is predicted to be in the α -helix, β -sheet, β -turn and coil conformation at each residue position, together with the corresponding graphical output. In the Robson prediction this gave also the averaged total score in units of centinats [defined by Garnier et al. (1978)].

The SP sequences found in six complement components were based on the following reports: factor D (White et al., 1992); factor I (Catterall et al., 1987); factor B (Mole et al., 1984); C2 (Bentley, 1986); C1r (Leytus et al., 1986; Journet and Tosi, 1986); C1s (Mackinnon et al., 1987). In the two-polypeptide proteins factor I, C1s and C1r, the SP sequence is initiated by the point of polypeptide cleavage that generated the light chain of factor I or the B-chain of C1r and C1s, i.e. starting at IIGG- or IVGG-. Sequences were aligned both manually and also with the use of guidelines provided by the automatic multiple sequence alignment program MULTAL (Taylor, 1990). Alignments were achieved by maximizing the occurrence of identical or chemically homologous residues in these sequences. Gaps were minimized in order to facilitate the averaging of structural properties. Molecular-graphics displays were analysed using Silicon Graphics Personal Iris 4D/25TG and Indigo XS24 Workstations with Insight II software (Biosym Technologies Inc.).

RESULTS AND DISCUSSION

Sequence analyses of SPs of known crystal structures

Figure 1 summarizes the topological sequence alignment of nine serine proteinases of known crystal structure. This was obtained

by the collation of results published in several crystallographic studies (see the Materials and methods section). Sequence similarity between the SP domains of known crystal structures can be readily rationalized in terms of the basic folding of the SP domain as two subdomains of predominantly β -sheet structure (Figure 2; Chothia and Janin, 1982). Each subdomain contains a six-stranded β -sheet barrel which can be described as a four-stranded Greek-key structural motif (β -strands B, C, D and F and β -strands J, K, L and M), which is followed by an antiparallel hairpin motif (β -strands G and H and β -strands N and O respectively). There are 12 key β -strands in all.

The main features of the alignment can be summarized as follows.

(i) Insertions and deletions

If the sequence alignment is made on the basis of the reported assignment of 15 β -strands A–O by the method of Kabsch and Sander (1983) in Read and James (1988), Figure 1 shows that the 230 principal residue positions (i.e. those occupied in over half of the nine sequences) are readily aligned as about ten contiguous segments of between 6 and 30 residues in length. There are only minor exceptions of single residue insertions or deletions. These ten segments occur at residues 16–36, 38–60, 64–95, 96–125, 128–146, 148–170, 171–177, 178–187, 189–202 and 207–245 (chymotrypsinogen numbering). Two larger deletions occur between β -strands H and I in *S. griseus* trypsin and between β -strands K and L in human elastase. Except for the short β -strand E, 14 of the 15 β -strands A–O are located within these ten segments. The positions of sequences of variable lengths between 1 and 11 residues between these segments are reproducibly located in the nine sequences. Inspection of the crystal structure shows that these variable segments are solvent-exposed polypeptide loops with no obvious constraint on their maximum sizes (Greer, 1990).

(ii) Residue conservation

The catalytic residues His⁵⁷-Asp¹⁰²-Ser¹⁹⁵ are totally conserved in well-defined structural segments in all nine sequences. Three other residues adjacent to His⁵⁷ (AAHC) and five more adjacent to Ser¹⁹⁵ (GDSGGP) are also totally conserved. In the latter, Asp¹⁹⁴ is responsible for the internal ion-pair formed with the N-terminus amino group. These are the most conserved regions in the nine sequences. More generally, among the 230 residues of the SP domain, 61 (27%) are conserved to better than 75% in the nine sequences, whereas 133 (57%) are conserved to better than 50%.

(iii) Predicted secondary structure

The comparison of the secondary structure observed in the crystal structures with that predicted as a consensus of the nine sequences is shown in Figure 1. In the averaged β -biased Robson

Figure 1 Comparative sequence alignment of nine crystal structures for the SP domain with the six SP domains found in the human complement components

The source of the sequences and abbreviations are given in the Materials and methods section. The chymotrypsinogen numbering is used; the dashes identify the residue positions used for secondary-structure averaging (i.e. occupied in five or more of the nine crystal structure sequences). Beneath each set of sequences, the disulphide bridge connectivity is denoted by pairs of numbers. That of C1s is reported by Hess et al. (1991). Potential oligosaccharide sites are underlined and are marked by 'c'. The catalytic triad of residues His⁵⁷-Asp¹⁰²-Ser¹⁹⁵ is denoted by hashes (#). The averaged Robson and Chou–Fasman predictions [see the Results and discussion section (iii)] and the Kabsch and Sander (K&S) secondary-structure assignments are denoted by a (α -helix), b (β -strand), 3 (3_{10} -helix) and t (turn). The 15 β -strands are lettered A–O to follow Table 9 in Read and James (1988). The six strands B, C, D, F, G and H and the corresponding six strands J, K, L, M, N and O are packed orthogonally in each of the two subdomains of the SP domain to form a β -barrel (Chothia and Janin, 1982). Above the complement sequences, the extent of residue conservation of a given class is specified by one or two * if this is better than 50 or 75% respectively in all 15 sequences. The C-terminal five to eight residues of factor B and C2 have been deleted for reasons of clarity. Abbreviations used: Bov, bovine; SG, *Streptomyces griseus*; Por, porcine; Hum, human.

Figure 2 Molecular-graphics views of the structure of β -trypsin

The two subdomains 1 and 2 are separated by the broken line. Flanking this broken line is the catalytic triad of residues His⁵⁷-Asp¹⁰²-Ser¹⁹⁵ (H, D and S in squares). Also indicated are the N-terminus and C-terminus (N and C), glycosylation sites (●) and disulphide bridges (○). These follow the SP sequence alignment shown in Figure 1. The 12 key β -strands B, C, D, F, G, H and J, K, L, M, N, O are indicated by arrowed ribbon outlines and the corresponding letter. (a) and (b) summarize respectively the structural data from nine crystallographically-determined SP sequences and those presumed for the six complement SP sequences.

prediction, the 12 key β -strands were all identified as clear peaks of high probability in the averaged total score of probabilities for α -helix, β -strand, turn and coil, with the partial exception of β -strand C. The averaged Chou-Fasman prediction located the 12 key β -strands as clear peaks, but with the exceptions of parts of β -strands C, G, H, J, L, M, N and O. The β -biased Robson prediction for this protein folding was thus the most successful. It should be noted that generally the typical level of agreement is around 50–60% accuracy in a full prediction of this type (Argos, 1989); the main source of error lies in the prediction of the start and end of the secondary-structure feature rather than in the prediction itself.

(iv) Disulphide bridges

All seven disulphide bridges (numbered 1, 2, 3, 6, 7, 8 and 9 in Figure 1) occur in well-defined identical positions within the ten contiguous segments (Figure 2a). The intrasubdomain disulphide bridges 3, 7, 8 and 9 are found in almost all the sequences. The less frequently occurring disulphide bridges 1, 2 and 6 connect the two subdomains of the SP domain.

(v) Glycosylation sites

Further inspection of the alignment shows that 11 of 12 potential glycosylation sites occur between the β -strands A to O, the twelfth one being on β -strand K. All lie in solvent-exposed locations as expected.

Analyses of the SP sequences in complement

Comparison of this sequence alignment and analysis with the sequences of the six SP domains found in the complement

proteins shows that all six sequences can be readily identified with structural features found in the crystal structures of SP proteins. Results are numbered to follow the above observations.

(i) Insertions and deletions

All but one of the ten sequence segments can be identified in the complement sequences. One exception is several three-to-nine-residue deletions found between β -strands F and G in factor B, factor I, C2 and C1r, in a region which is, however, seen from molecular-graphics inspection to be on the protein surface, where excisions can be readily achieved. There is a two-residue insertion in the factor I light-chain sequence near to disulphide bridge 1, i.e. close to Cys¹²⁴ (chymotrypsinogen numbering; Figure 1) which is linked with the heavy chain of factor I. The location of variable lengths of sequences outside the ten segments occurs in similar positions in the complement SP domains. Some notable increases in the size of several insertions are observed for factor B, C2, C1r and C1s, such as those between residues 119 and 120, 146 and 148 and 177 and 178. For factor B, C1r and C1s, Caporale et al. (1984) and Fothergill et al. (1989) have suggested that these changes correspond to surface insertions and deletions at the protein surface close to the catalytic site. In justification of these premises, Figure 1 shows that none of these insertions or deletions occurs in core β -strands of the β -sheet structure.

(ii) Residue conservation

Residue conservation is highest in the regions marked by double asterisks in Figure 1, and applies to only 59 (26%) out of the 230 residues in the standard SP sequence. These conserved regions correspond largely to the 12 key β -strands. The three catalytic

residues His⁵⁷-Asp¹⁰²-Ser¹⁹⁵ are fully conserved in the expected positions, together with almost all the adjacent totally conserved residues noted above (AAHC and GDSGGP). One exception is the relatively minor disappearance of Cys⁵⁸ in C1r and C1s; the other is the conservative replacement of Asp¹⁹⁴ in C2 with Glu¹⁹⁴. Gagnon (1984), Mole et al. (1984) and Greer (1990) have, however, noted that, for C2 and factor B, the N-terminal sequence from residue 16 onwards is atypical, the sequences CGG at residues 42–44 (β -strand C) and G-aromatic-G at residues 140–142 (β -strand J) are not conserved, and difficulties in sequence alignment between residues 127 to 167 are observed. These indicate localized structural differences in both the SP subdomains (Figure 2b), but their significance is not clear.

(iii) Predicted secondary structure

For each of the six complement SP sequences, the β -biased Robson prediction (results not shown) was compared with both the averaged prediction from the nine crystallographic sequences and the observed location of the β -strands. This predictive method had been the most successful for the SP sequences above. Despite the use of non-averaged predictions, the 12 key β -strands that define the SP fold were identified wholly or mostly as β -strands in each of the six sequences. The only complete exceptions were β -strand D in factor D, β -strand H in factor B and C2, and β -strand G in C1r. This general agreement for 11 or 12 β -strands suggests that the basic β -sheet structure is retained in all six complement SP domains.

(iv) Disulphide bridges

Disulphide bridges can be readily identified at all the positions numbered 1, 3, 7, 8 and 9, as observed in the SP crystal structures, but not at positions 2 or 6. The intersubdomain disulphide bridge 1 of α -chymotrypsin now connects the A and B chains of C1r and C1s and the light and heavy chains of factor I. Two non-homologous disulphide bridges 4 and 5 are found. One of these occurs at position 4 in factor I between Cys⁵⁰ and Cys¹¹⁷ in subdomain 1. Inspection of the crystal structure of β -trypsin shows that the α -carbon atoms are too far apart at 1.6 nm (instead of the desired distance of 0.5 to 0.6 nm). This disulphide bridge can only be formed if structural rearrangements corresponding to the two-residue insertion have occurred (50–117 in Figure 2b). The other of these at position 5 occurs only in C2 and factor B, and occurs in the region that cannot be aligned with other SP sequences in subdomain 2. The α -carbon atoms of the putative Cys¹²⁷ and Cys¹³⁹ residues are separated by 2.2 nm (127–139 in Figure 2b). This implies that a structural rearrangement is required before these can be connected, and the large adjacent insertion between residues 146 and 148 may be involved in this. While five disulphide bridges in the complement SP sequences are fully consistent with the SP crystal structures, two more are atypical and involve at least a localized structural rearrangement before they can be formed.

(v) Glycosylation sites

Ten putative glycosylation sites are found in four of the six SP sequences. The exceptions are factors B and D. All ten occur outside the region expected to be β -strands. Of these ten sites, six are found in loop regions of various sequence lengths. Eight of the ten sites are in regions expected to be solvent-exposed, which

is consistent with the protein folding of the SP structure. In the crystal structure (Figure 2b), all ten sites are distant from the catalytic triad His⁵⁷-Asp¹⁰²-Ser¹⁹⁵ and cannot obstruct access to this, as desired. The remaining two putative sites occur near residue 16 of C2 in the region expected to form the N-terminal salt-bridge with Asp/Glu¹⁹⁴ (Figure 2b); thus a localized conformational rearrangement is expected in this region.

These five considerations show that all six SP domains in the complement proteins are structurally homologous with the basic SP crystal structure. Several instances occur where large localized differences are seen for C2, factor B and factor I. All of these, nonetheless, appear to be compatible with the core SP structure identified from crystallography. For factor D, the recent determination of its crystal structure shows that structural homology is indeed the case (Narayana et al., 1991a,b).

CONCLUSIONS

The percentage identity of the α -chymotrypsin sequence with those for the SP domain in factor D, factor I, factor B, C2, C1r and C1s is low at 43, 41, 33, 31, 35 and 42% for conservatively replaced residues. These similarities are to be compared with the range of 27–57% for the degree of conservation found in the crystal structure sequences of the SP fold. A sequence alignment that is compatible with the structural features in nine crystal structures should be adequate evidence that the sequence itself corresponds to a homologous protein structure, despite the observed differences in biochemical properties and gene structures. Here the alignment of six complement SP sequences with the average of nine SP crystal structures has been justified on the grounds of conserved residues and disulphide bridges, solvent-exposed glycosylation sites, together with the demonstration that sequence insertions and deletions can be located on the protein surface, and the general consistency with secondary-structure predictions.

This identification facilitates analyses of structural data on the complement SP domains. In the modelling of solution scattering data for factor D, only minor alterations were required to the structure of β -trypsin to model the scattering curves for factor D [the accompanying paper (Perkins et al., 1993a)], and likewise to factor I where more significant alterations were needed [the other accompanying paper (Perkins et al., 1993b)]. The β -biased Robson secondary-structure prediction from this work gives the β -sheet contents of factor D, factor I, factor B, C2, C1r and C1s as 52–64% and their α -helix contents as 2–11%. These predictions can be compared with quantitative Fourier-transform i.r. spectroscopy data on the secondary structures of these proteins (S. J. Perkins, K. F. Smith, S. C. Williams, P. I. Haris, D. Chapman and R. B. Sun, unpublished work). Factor B, C2 and factor I, however, exhibit several localized differences in the SP domain. This will affect the success of molecular-graphics modelling in attempting to predict an accurate atomic model from its sequence homology. In addition, the existence of large loop insertions in C1r, C1s, factor B and C2 make these segments difficult to model, as these have many more degrees of conformational freedom (Greer, 1990).

We thank the Wellcome Trust and The Royal Society for their support, Mr. A. S. Nealis for computational assistance, and Dr. R. B. Sim for carefully reading this manuscript and for useful discussions.

REFERENCES

- An-Zhi, W., Mayr, I. and Bode, W. (1988) FEBS Lett. **234**, 367–373
- Argos, P. (1989) in Protein Structure: A Practical Approach (Creighton, T. E., ed.), pp. 169–189, IRL Press, Oxford

- Bentley, D. R. (1986) *Biochem. J.* **239**, 339–345
- Bode, W. and Schwager, P. (1975) *J. Mol. Biol.* **98**, 693–717
- Bode, W., Chen, Z., Bartels, K., Kutzbach, C., Schmidt-Kastner, G. and Bartunik, H. (1983) *J. Mol. Biol.* **164**, 237–282
- Campbell, R. D. and Porter, R. R. (1983) *Proc. Natl. Acad. Sci., U.S.A.* **80**, 4464–4468
- Caporale, L. H., Woods, D., Gagnon, J., Christie, D. L. and Bing, D. H. (1984) *Immunobiology* **164**, 213 (abstr.)
- Catterall, C. F., Lyons, A., Sim, R. B., Day, A. J. and Harris, T. J. R. (1987) *Biochem. J.* **242**, 849–856
- Chothia, C. and Janin, J. (1982) *Biochemistry* **21**, 3955–3965
- Chou, P. Y. and Fasman, G. D. (1978) *Adv. Enzymol. Relat. Areas Mol. Biol.* **47**, 45–148
- Cohen, G. H., Silverton, E. W. and Davies, D. R. (1981) *J. Mol. Biol.* **148**, 449–479
- Fothergill, J. E., Kemp, G., Paton, N., Carter, P. and Gray, P. (1989) *Behring Inst. Mitt.* **84**, 72–79
- Fujinaga, M. and James, M. N. G. (1987) *J. Mol. Biol.* **195**, 373–396
- Gagnon, J. (1984) *Philos. Trans. R. Soc. London B* **306**, 301–309
- Garnier, J. and Robson, B. (1989) in *Prediction of Protein Structure and the Principles of Protein Conformation* (Fasman, G. D., ed.), pp. 417–465, Plenum Press, New York
- Garnier, J., Osguthorpe, D. J. and Robson, B. (1978) *J. Mol. Biol.* **120**, 97–120
- Gibrat, J.-F., Garnier, J. and Robson, B. (1987) *J. Mol. Biol.* **198**, 425–443
- Greer, J. (1990) *Protein: Struct. Funct. Genet.* **7**, 317–334
- Hess, D., Schaller, J. and Rickli, E. E. (1991) *Biochemistry* **30**, 2827–2833
- Ishikawa, N., Nonaka, M., Wetsel, R. A. and Colten, H. R. (1990) *J. Biol. Chem.* **265**, 19040–19046
- James, M. N. G., Delbaere, L. T. J. and Brayer, G. D. (1978) *Can. J. Biochem.* **56**, 396–402
- Journet, A. and Tosi, M. (1986) *Biochem. J.* **240**, 783–787
- Kabsch, W. and Sander, C. (1983) *Biopolymers* **22**, 2577–2637
- Law, S. K. A. and Reid, K. B. M. (1988) *Complement*, IRL Press, Oxford
- Lewin, R. (1987) *Science* **237**, 1570
- Leytus, S. P., Kurachi K., Sakariassen K. S. and Davie, E. W. (1986) *Biochemistry* **25**, 4855–4863
- Mackinnon, C. M., Carter, P. E., Smyth, S. J., Dunbar, B. and Fothergill, J. E. (1987) *Eur. J. Biochem.* **169**, 547–553
- Mole, J. E., Anderson, J. K., Davison, E. A. and Woods, D. E. (1984) *J. Biol. Chem.* **259**, 3407–3412
- Narayana, S. V. L., Kilpatrick, J. M., El-Kabbani, O., Babu, Y. S., Bugg, C. E., Volanakis, J. E. and DeLucas, L. J. (1991a) *J. Mol. Biol.* **219**, 1–3
- Narayana, S. V. L., Carson, M., DeLucas, L., Moore, D., Bugg, C. E., Kilpatrick, J. M. and Volanakis, J. E. (1991b) *Complement Inflamm.* **8**, 198 (abstr.)
- Navia, M. A., McKeever, B. M., Springer, J. P., Lin, T.-Y., Williams, H. R., Fluder, E. M., Dorn, C. P. and Hoogsteen, K. (1989) *Proc. Natl. Acad. Sci., U.S.A.* **86**, 7–11
- Perkins, S. J., Haris, P. I., Sim, R. B. and Chapman, D. (1988) *Biochemistry* **27**, 4004–4012
- Perkins, S. J., Smith, K. F., Kilpatrick, J. M., Volanakis, J. E. and Sim, R. B. (1993a) *Biochem. J.* **295**, 87–99
- Perkins, S. J., Smith, K. F. and Sim, R. B. (1993b) *Biochem. J.* **295**, 101–108
- Read, R. J. and James, M. N. G. (1988) *J. Mol. Biol.* **200**, 523–551
- Reeck, G. R., de Haën, C., Teller, D. C., Doolittle, R. F., Fitch, W. M., Dickerson, R. E., Chambon, P., McLachlan, A. D., Margoliash, E., Jukes, T. H. and Zuckerkandl, E. (1987) *Cell* **50**, 667
- Reid, K. B. M. (1986) *Essays in Biochemistry* **27**, 27–68
- Reid, K. B. M., Bentley, D. R., Campbell, R. D., Chung, L. P., Sim, R. B., Kristensen, T. and Tack, B. F. (1986) *Immunol. Today* **7**, 230–234
- Remington, S. J., Woodbury, R. G., Reynolds, R. A., Matthews, B. W. and Neurath, H. (1988) *Biochemistry* **27**, 8097–8105
- Rogers, J. (1985) *Nature (London)* **315**, 458–459
- Sali, A. and Blundell, T. L. (1990) *J. Mol. Biol.* **212**, 403–428
- Sawyer, L., Shotton, D. M., Campbell, J. W., Wendell, P. L., Muirhead, H., Watson, H. C., Diamond, R. and Ladner, R. C. (1978) *J. Mol. Biol.* **118**, 137–208
- Sim, R. B. and Perkins, S. J. (1989) *Curr. Top. Microbiol. Immunol.* **153**, 209–222
- Sprang, S., Standing, R. J., Fletterick, R. J., Stroud, R. M., Finer-Moore, J., Xuong, N.-H., Hamlin, R., Rutter, W. J. and Craik, C. S. (1987) *Science* **237**, 905–913
- Taylor, W. R. (1990) *Methods Enzymol.* **183**, 456–474
- Tosi, M., Duponchel, C., Meo, T. and Couture-Tosi, E. (1989) *J. Mol. Biol.* **208**, 709–714
- Tsukada, H. and Blow, D. M. (1985) *J. Mol. Biol.* **184**, 703–711
- Wang, D., Bode, W. and Huber, R. (1985) *J. Mol. Biol.* **185**, 595–624
- White, R. T., Damm, D., Hancock, N., Rosen, B. S., Lowell, B. B., Usher, P., Flier, J. S. and Speigelman, B. M. (1992) *J. Biol. Chem.* **267**, 9210–9213