An organ system without a proper name

As every high-school graduate knows, animals are classified into two large groups – those that have vertebrae and those that do not. What every high-school graduate should also know (but probably doesn’t) is that the vertebrae are further divided into jawless (Agnatha) and jawed (Gnathostomata). The jawed vertebrates are distinguished not only by the possession of jaws, but also by a suite of other characters, including an immune system based on the use of the “holy trinity” or three sets of molecules – the major histocompatibility complex (MHC) molecules, the T-cell receptors (TCRs), and the B-cell receptors or immunoglobulins (BCRs or Igs; see Klein et al., 2000). The system is referred to as adaptive or acquired, but both designations are misleading and inaccurate; a term like anticipatory immune system (Klein, 1989) would have, perhaps, been more appropriate, but it did not take. Fortunately, the acronyms for all three names are the same – AIS. Immunologists originally believed that the AIS is present in both vertebrates and nonvertebrates (Cooper, 1976), but I have argued for some time (Klein, 1989, 1997) that the system is restricted to vertebrates and possibly to jawed vertebrates only. The first part of this proposition is now generally accepted; the second part is still debated, but this debate, too, is close to resolution. If we accept the holy trinity as the defining feature of the AIS, mounting evidence indicates that the agnathans lack the system, for the trinity appears to be present in all gnathostomes, but absent in jawless vertebrates.

The most compelling evidence for the absence of the trinity in jawless vertebrates has been provided by recent studies in which cells resembling mammalian lymphocytes have been purified from lamprey organs and cDNAs obtained from the gene transcripts present in these cells have been sequenced (Mayer et al., 2002a; Uinuk-Ool et al., 2003 and unpublished data). By this approach, some 30,000 lamprey cDNA sequences have been collected and through them numerous genes expressed in the lymphocyte-like cells have been identified. However, no members of the trinity could be found in the collection. Although definitive evidence for the absence of the trinity will have to await the completion of the entire lamprey genome sequence, the cDNA sequences suggest strongly that MHC, TCR, and BCR genes as we know them from gnathostomes are absent in the lamprey and probably also in the hagfish. Assuming this to be the case, an inquiry into the nature of the agnathan immune system can be expected to yield valuable information about the transition from non-AIS to AIS at the time of divergence of jawless and jawed vertebrates.

Origin of the holy trinity

In its fully developed mammalian form the AIS is tremendously complex, encompassing directly four organs (thymus, bone marrow together with the bursa of Fabricius in birds, spleen, and lymph nodes with Peyers’ patches) and indirectly nearly all organs except the immunologically privileged ones (Klein and Hoefijz, 1997). Although centered on lymphocytes and antigen-presenting cells such as dendritic cells and macrophages, the AIS is contingent on the participation of a large number of other cell types. And although three groups of molecules (the holy trinity) can be singled out as being paradigmatic of the system, the actual number of molecules without which the system becomes seriously handicapped probably goes into thousands. An appreciable portion of the transcribed human genome is directly or indirectly concerned with body’s defence against parasites of all kinds.

The deciphering of the origin of such a complex system will therefore keep generations of immunologists and evolutionary biologists busy. At the top of the list of priorities in this effort is the origin of the holy trinity. The problem of the origin of the MHC molecules can be decomposed into two – the derivation of the peptide-binding domain and the ancestry of the immunoglobulin-like domains, together with the transmembrane and cytoplasmic regions (Klein and O’hUigin, 1993). There is not much to say about the former, for no relative of this
domain has been identified outside the group of jawed vertebrates. It is therefore not very productive to speculate about its origin, except for treating any peptide-binding protein as a potential candidate. We are in a somewhat better position in regard to the immunoglobulin-like domain of the MHC molecules and the immunoglobulin domains of the TCRs and BCRs, which all seem to be related to one another (Kasahara et al., 2004). The domains are distinguished into two types, V (variable) and C1 (constant); several other types occur in some of the other members of the very large immunoglobulin superfamily (Halaby et al., 1999). The types differ in the length and arrangement of the \( \beta \)-strands in and between the two \( \beta \)-pleated sheets of the sandwich-like tertiary structure of the Ig domain. The MHC molecules contain the C1-type domains, which are also present in the TCR and BCR molecules, as well as in several other vertebrate proteins such as tapasin (possibly an MHC-derivative, see Mayer and Klein, 2001), butyrophilin, and signal-regulation proteins (SIRPs). The C1-type has thus far been found only in jawed vertebrates (Kasahara et al., 2004). The TCR and BCR molecules contain in addition to the C1 also the V-type domain, which is known to be present in several other vertebrate molecules (Kasahara et al., 2004), as well as in the cephalochordate Amphioxus (Cannon et al., 2002; Sato et al., 2003) and the urochordate Ciona (Azumi et al., 2004) molecules. The other types of Ig domain are widely distributed among the various eukaryotic and even prokaryotic members of the Ig superfamily. The domains are capable of binding a variety of different ligands and so participate in cell adhesion, cell-cell interaction, and polymerization of polypeptide chains. Thus, although in this case a direct ancestor of the Ig domains in the MHC, TCR, and BCR molecules cannot be identified either, molecules not too far removed from it are known. Since the various ancestral molecules apparently existed before the actual emergence of the AIS, they probably did not function as antigen receptors and their encoding genes did not rearrange somatically. Structurally, however, they might have been nearly ready to assume these functions.

**Origin of the thymus**

Of all the organs involved in immunity, the thymus appears to be the *sine qua non* of the AIS. All the other organs are either present in only some jawed vertebrates (e.g., lymph nodes) or their primary (original?) function is not concerned with immunity directly (e.g., spleen). Morphologically differentiated thymus has been found in all studied jawed vertebrates. The claimed presence of a thymus in the lamprey was based on an observed accumulation of cells with lymphocyte morphology in the pharyngeal region (see Good et al., 1972; Ardavin and Zapata, 1988). Since, however, accumulation of cells alone does not make an organ and since lymphocytes tend to accumulate transiently or permanently along the entire digestive tract (Fichtelius, 1970), these observations cannot be accepted as evidence that lampreys possess a thymus. The issue remains unresolved and its resolution will require more than morphological investigation alone. In particular, the recent studies on molecular mechanisms of thymus determination in the mouse (Manley, 2000; Blackburn and Manley, 2004) need to be extended to jawed fishes and to jawless vertebrates.

In mouse embryos, outpocketings of the pharyngeal endoderm give rise to four pharyngeal pouches characterized by the expression of paired box genes 1 and 9 (*Pax1, Pax9*) and fibroblast growth factor 8 (*Fgf8*). Two of the pouches (p3 and p4) are distinguished from the other two (p1 and p2) by the expression of the homeobox A3 (*HoxA3*) gene. Subsequent interaction of the p3 and p4 endoderm with the ectodermally derived neural crest cells initiates the formation of a joined thymus-parathyroid gland rudiment. The process is controlled by the *HoxA3*, *Pax1/Pax9*, eye absent homologue 1 (*Eya1*), and sine oculis-related homeobox homologue 1 (*Six1*) gene, acting in this order in a cascade-like fashion. The rudiment then differentiates into two domains, one characterized by the expression of the glial cells missing homologue 2 (*Gem2*) and the other by the expression of the forkhead box N1 (*foxn1*) gene, the former giving rise to the parathyroid gland and the latter to the thymus. Subsequently, the rudiment detaches from the pharynx, migrates to the heart region, and the two domains separate from each other to assume their respective positions. The parathyroid gland becomes embedded in the thyroid and the thymus settles down on the pericardium, where it is invaded by thymocyte progenitor cells derived from lateral plate mesoderm. In contrast to tetrapods, fishes, including agnathans, have generally more pharyngeal arches and their pharyngeal pouches contribute mainly to the development of feeding and breathing apparatuses. The parathyroid gland is present in tetrapods only, but parathyroid hormones have been identified in both bony and cartilaginous fishes. The site of production of the hormones in these fishes has not been determined and no parathyroid hormones have been reported from agnathans. In jawed fishes, thymus rudiments arise in all of the numerous pharyngeal pouches, but the molecular mechanisms of their determination have not been elucidated. The transcription factors determining the thymus rudiment in mammals evolved long before the rise of vertebrates and so it seems probable that their homologues participate also in the development of fish thymi. Identification of their agnathan homologues should therefore offer an opportunity to study their expression in the pharyngeal region at the time of pouch formation in this vertebrate group and thus throw light on the origin of the thymus.

**Origin of lymphocytes**

While there is thus far no convincing evidence for the presence of the thymus in any of the living jawless vertebrates, there can now be little doubt that at least the lam-
preys possess lymphocyte-like cells. Light microscopists have long known about the presence of cells with typical lymphocyte morphology (small size, round contour, thin cytoplasmic rim, compacted nuclear heterochromatin) in not only agnathans, but also some nonvertebrate and non-chordate deuterostomes (Ratcliffe and Rowley, 1981). However, since morphological resemblance is an unreliable guide to cell identity, the presence of lymphocyte-like cells in agnathans remains uncertain. The doubts have now been dispelled by studies on purified populations of these cells (Mayer et al., 2002a; Uinuk-Ool et al., 2002). These studies not only confirm the findings of the light microscopists, but extend them by demonstrating that the similarity holds also at the electron microscopic level (Mayer et al., 2002a) and is further supported by a series of other findings. The lamprey cells have been shown to express homologues of genes known in mammals to be involved in the control of lymphopoiesis. They include the homologue of the spleen focus-forming virus integration B (Spi-B) gene (Shintani et al., 2000) and genes homologous to the mammalian “sun” genes (Ikaros, Helios, Aiolos, Eos; see Haire et al., 2000; Mayer et al., 2002b). Essential steps in the vertebrate lymphohaemopoietic pathway must have therefore been in place before the divergence of jawless and jawed vertebrate lineages and hence before the emergence of the AIS.

The strongest support for the lymphocyte-like character of the lamprey cells is provided by their gene expression profile (Uinuk-Ool et al., 2002), which is reminiscent of that of mammalian lymphocytes. The genes include the homologues of CD45, CD9/CD81, B-cell adaptor for phosphoinositide 3-kinase (BCAP), CD3ε-associated signal transducer (CAST), CD98, and transporter associated with antigen processing (TAP), which in mammals are involved in regulation of T- and B-cell stimulation and proliferation; stabilization of molecular complexes participation in lymphocyte activation, adhesion, migration, and differentiation; adaptor functions in signalling leading to the activation of B and T lymphocytes; amino acid transport associated with cell activation; and delivery of peptides to the MHC molecules; respectively. The CD45 homologue has also been demonstrated in the hagfish, another agnathan (Nagata et al., 2002). Despite all these resemblances it is probably prudent not to call the lamprey cells “lymphocytes”, for there is no evidence that they are able to undergo any kind of clonal expansion, which is one of the defining features of true lymphocytes (Klein, 1988); the designation “lymphocyte-like cells” is more appropriate at this stage.

Origin of antigen processing and presentation

Waste disposal is a problem which arises with product manufacture. An introduction of product making must therefore always go hand in hand with product degradation. Indeed, synthetic and degradation mechanisms are among the processes that evolved earliest in the evolution of life. The degradation mechanisms include phagocytosis, apoptosis, and proteasomal protein degradation, all of which have been appropriated by the AIS. Of the three, proteasomal protein degradation is most directly coupled with the initiation of the adaptive immune response. All of the numerous subunits, from which the molecular shredding machine called the “proteasome” is assembled, have a long evolutionary history which goes back to at least the origin of the eukaryotes. Of course, if it was to serve the needs of the AIS, minor adjustments in the molecular machine had to be made. So as not to jeopardize the main disposal mechanism, the adjustments were made on copies of the existing genes produced by the duplication of the genes encoding certain of the subunits. The adjustments were of two kinds. First, the specificity of the enzymatically active subunits was altered to ensure that the peptides coming out of the shredder were of the right length and composition to fit into the cleft of the MHC peptide-binding domain. And second, the synthesis of these subunits and the construction of specialized “immunoproteasomes” were brought under the control of factors such as interferon gamma, released during an immune response. Both types of adjustment were apparently achieved by a small number of substitutions in the protein-encoding and regulatory regions of the gene. It appears that the lamprey possesses genes for the subunits before these substitutions occurred in them; they must have occurred in the gnathostome lineage after its separation from the agnathan lineage (Takezaki et al., 2002, 2003).

Minor adjustments were also necessary in the transporter proteins which ferry peptides emerging from the proteasomal shredder across the membrane of the endoplasmic reticulum for loading onto the MHC molecules. Here again, the transporter-encoding gene present in the lamprey appears to be waiting for changes to occur which focus its specificity on peptides accomodable by the MHC molecules on the luminal side of the membrane (Uinuk-Ool et al., 2003).

The origin of the GOD

An essential feature of the AIS is the somatic diversification of the TCRs and BCRs in clones of proliferating T and B lymphocytes, respectively. Although initially some diversity of the receptors might have been generated by multiplication and diversification of their encoding genes in the germ line, its extent must have been limited by necessity. Ultimately, a diversification mechanism evolved which was unlimited in range, dynamic, and open to continuous upgrading – a somatic generator of diversity or GOD. This evolution started from a membrane-bound protein expressed in the lymphocyte-like cells – the ancestor of all the TCRs and BCRs. It consisted of two extracellular Ig domains, one V and one C1, in addition to the transmembrane and
cytoplasmic regions. It was encoded in a single gene in which the V and C1 domain were specified by separate exons. The V exon subsequently split into two segments, a longer V and shorter J, separated by an intervening sequence. The gene then duplicated repeatedly and in some of the copies the V exon split off an even shorter D segment which, however, will largely be ignored in further discussion for the sake of simplicity. Whatever the mechanism of the splitting was, it left all the segments flanked by tripartite signal sequences, the V segment at one end and the J segment at the opposite end. Further multiplication of the V and J segments, followed by their mutational diversification, created a germline found of receptor diversity to be expanded somatically by the GOD in lymphocyte clones.

The essence of the GOD is to bring together randomly one V and one J segment and to express each pair together with the same C exon separately in a different lymphocyte clone. The pairing of the segments is achieved by a process resembling recombination in which two genes, the recombination activating genes 1 and 2 (RAG1 and RAG2), play a leading role. By recognizing the signal sequences flanking the V and J segments and binding to them, the proteins encoded in these genes bring always together one V and one J segment, and with the help of other proteins excise the entire sequence separating the chosen segments and join these together. Imprecision in the cutting and pasting together of the segments further diversifies the resulting products. This V-J recombination process is strikingly similar to the mechanism by which transposons move from one place to another in a genome and between genomes. Specifically, the RAG1 and RAG2 proteins can be viewed as modified transposases (enzymes intimately associated with the movement of transposons), which bind to the signal sequences, thus initiating a reaction that excises the sequence between the two chosen segments without, however, integrating it (as a true transposon would) elsewhere in the genome. Indeed, in vitro tests have demonstrated that the RAG proteins are still capable of carrying out this integration function which they no longer do in vivo, presumably being inhibited from doing so to avoid messing up the genome (Agrawal et al., 1998; Hiom et al., 1998).

These observations have led to the hypothesis that the RAG genes and the signal sequences flanking the V, J, and D segments were introduced into the genome of an ancestral jawed vertebrate by a transposon originating in a bacterium or virus (Bernstein et al., 1996; Agrawal et al., 1998; Hiom et al., 1998).

According to this hypothesis, the original transposon consisted of the two RAG genes in their in vivo-active form, flanked by the signal sequences. After its integration into the genome of the ancestral vertebrate, it became active and produced a variant lacking the RAG genes, which then transposed once or twice into the V-exon of the primigenial receptor gene, splitting it into the V and J segments flanked by the signal sequences. (The second transposition may have generated the D segment which, unlike the V and J segments, is flanked by signal sequences at both its ends.) The original integrated transposon then lost its flanking signal sequences, as well as the ability to transpose itself in vivo. Instead, the two RAG genes became involved exclusively in the V (D) J recombination of the emerging TCR and BCR genes.

In addition to the evidence cited above, two other observations support the transposon hypothesis. First, low sequence similarity of the RAG1 protein with bacterial and viral transposases has been reported (Bernstein et al., 1996). And second, the two RAG genes, although only distantly related to each other at best, are closely juxtaposed in all the species studied and are transcribed concomitantly, as might be expected of genes in a compact transposon genome. On the other hand, although the resemblances to a transposase are striking, they do not exclude alternative interpretations. Hughes (1999), for example, has pointed out that the RAG1 protein also shows sequence similarity to the yeast RAD18 and to related fungal DNA-binding proteins, while the RAG2 amino acid sequence is similar to SAS, another eukaryotic DNA-binding protein. He suggests two alternative interpretations of the similarities. First, instead of transposons giving rise to RAGs, the transposons themselves could have originated from recombination proteins whose genes might have escaped from the vertebrate genome. And second, the “transposase-like features of RAG1 and RAG2 have simply arisen as a fortuitous by-product of these molecules’ function in segmental rearrangement” (Hughes, 1999, p. 352).

**Big Bang or gradual evolution?**

We live in an era of superlatives. In science, we have superfamilies, superantigens, novelties (rather than new things) left and right, and Big Bangs or BBs. To be sure, the cosmical BB was big and it was a bang, lasting but a fraction of a second. Now, however, some paleontologists see one BB following another in the history of life, the biggest of them all at the onset of the Cambrian – the “Cambrian explosion”. Some immunologists, too, came up with their BB – the origin of the AIS (Bernstein et al., 1996). To a skeptical mind, these biological BBs seem neither big nor sudden when viewed in the right time frame and from the right perspective. From the point of view of molecular evolution, the time frames of both the Cambrian explosion and the emergence of AIS seem to be getting longer and longer. An “explosion” that lasts some 100 million years or longer surely seems more like a firecracker that failed to burst.

I agree with Hughes (1999) that calling the emergence of AIS a BB is, to put it mildly, an exaggeration. The evidence emerging from the study of the agnathans clearly points in the direction of small changes accumulating
over long periods of time, rather than to any sudden upheavals. Come to think of it, how could it have been otherwise in a system with so many different components which all had to be integrated and harmonized with one another for the whole to work? In many respects, the lampreys come quite close to possessing an AIS; all that is missing are the few final pieces to give the whole system another dimension. Even if transposons of viral or bacterial origin did contribute in the final stage of the emergence of AIS, which is far from certain at this stage, it was certainly not a one-night affair. More likely, a long period of functional integrative adjustments must have followed the initial event. Somehow, I am unable to hear the supposed bang – but then I might just be getting too old for such a novel event.

Jan to Jan

Although Professor Jan Svoboda (whose seventieth birthday this Festschrift celebrates) and I shared a mentor, Professor Milan Haöek, our scientific interests diverged early on in our careers. The present essay is a meek attempt at converging them. In reality, however, through all these years, we have shared more than the same first name. Like the quantum particles which, though separated in space and time after launching, somehow know what the other is doing, we have remained in touch. From one particle to the other the message reads: Happy Birthday, Jan, and – as they say, many happy returns.

Acknowledgements

I thank Kathleen Seasholtz for help in the preparation of the manuscript.

References


Sato, A., Uimuk-Ool, T., Kuroda, N., Mayer, W. E., Takezaki, N., Dongak, R., Figueroa, F., Cooper, M. D., Klein, J.


