The origin and formation of clay minerals in soils: past, present and future perspectives

M. J. WILSON

Macaulay Land Use Research Institute, Craigiebuckler, Aberdeen AB15 8QH, UK

(Received 23 September 1997; revised 15 January 1998)

ABSTRACT: The origin and formation of soil clay minerals, namely micas, vermiculites, smectites, chlorites and interlayered minerals, interstratified minerals and kaolin minerals, are broadly reviewed in the context of research over the past half century. In particular, the pioneer overviews of Millot, Pedro and Duchaufour in France and of Jackson in the USA, are considered in the light of selected examples from the huge volume of work that has since taken place on this topic. It is concluded that these early overviews may still be regarded as being generally valid, although it may be that too much emphasis has been placed upon transformation mechanisms and not enough upon neoformation processes. This review also highlights some of the many problems pertaining to the origin and formation of soil clays that remain to be resolved.

It has long been recognized that the minerals in the clay (<2 µm) fractions of soils play a crucial role in determining their major physical and chemical properties, and inevitably, questions concerning the origin and formation of these minerals have assumed some prominence in soil science research. This review considers some important aspects of these questions and is confined to clay minerals in soils as they are generally understood, that is as a medium for plant growth.

The occurrence of clay minerals in saprolites or weathered rock is not discussed to any great extent because many studies have shown that saprolites are characterized by physicochemical conditions that are different to those in soils. Intimate grain-to-grain contacts promote a special chemical environment on a local scale, bringing about the formation of transient clay mineral phases which quickly disappear in the overlying soil.

For the purposes of this review the perspective of the past will be taken to be that of the early 1960s, which is a convenient historical benchmark in that it represents a time when the overall structural characteristics of the layer silicate clay minerals were to a large extent known, even though much detail remained to be filled in, as well as a time that immediately pre-dated the widespread utilization in soil science of analytical techniques such as scanning electron microscopy, electron probe microanalysis, Mössbauer spectroscopy, electron spin resonance spectroscopy and infrared spectroscopy. This review will therefore attempt to summarize the general conclusions that had been arrived at regarding the origin and formation of certain clay minerals in soils in the early 1960s, the developments since that time and the situation as it currently stands, and to consider the outstanding problems to be addressed in the future. In this context, particular attention will be paid to micas, vermiculites, smectites, chlorites, interstratified minerals and kaolin minerals. As in all such reviews, a certain amount of subjectivity and personal preference is perhaps inevitable.

PAST PERSPECTIVES

The reviews of Millot (1965) and Jackson (1964) represent a convenient starting point with regard to the origin and formation of soil clays as assessed principally by the X-ray diffraction and microscopic
techniques that were available at the time. Millot distinguished three principal processes to account for the genesis of clay minerals, which may occur at different points in the geochemical cycle including weathering or soil formation at the earth’s surface. These processes are: (a) detrital inheritance whereby, for soils, clay minerals are inherited from pre-existing parent rock or weathered materials; (b) transformation where the essential silicate structure of the clay mineral is maintained to a large extent, but with major change in the interlayer region of the structure; and (c) neoformation, where the clay mineral forms through crystallization of gels or solutions.

Inherited soil clays may be of an extremely diverse and complex nature, reflecting both the variety of the parent rock as well as the transformation and neoformation processes that may have occurred in previous weathering environments. In order to understand current clay mineral-forming processes in a soil, it is essential that the contribution of inheritance is clearly understood.

With regard to transformation of clay minerals, a general example would be:

\[ \text{illite} \rightarrow \text{vermiculite} \rightarrow \text{smectite}. \]

This reaction proceeds through a process of depletion and exchange of interlayer K and concomitant decrease of layer charge. Such changes are, however, deceptively simple and have given rise to much debate about the precise mechanisms involved, as will be discussed later. Millot distinguished ‘degradation’ and ‘aggradation’ as separate forms of the transformation process. The above conversion of illite to montmorillonite involves depletion of elements from illite and is termed degradation, but the reverse reaction (aggradation) involves addition of K and other elements. Millot considered that degradation was characteristic of weathering rocks and soils, but that aggradation was rare in such environments.

The formation of clay minerals through neoformation clearly depends upon the appropriate physicochemical conditions of the immediate weathering environment, such as the pH, composition and concentration of the soil solutions, as well as the nature of the starting material and factors relating to the external environment like temperature, rainfall and percolation rate. Millot described kaolinite and montmorillonite as classical products of neoformation in soils forming under contrasting conditions. Thus, kaolinite is typical of freely drained, acid and base-depleted tropical environments, where an abundant supply of water ensures the required silica and alumina. Montmorillonite is more typical of a poorly drained or hydromorphic soils under alkaline conditions, rich in Mg and Ca ions and where Si, Al and commonly Fe tend to accumulate.

These generalizations were further developed and synthesized by French soil scientists, most notably Pedro and Duchaufour. Pedro (1964) distinguished two mineral weathering processes implicated in the formation of clay minerals in soils which he termed ‘hydrolysis’ and ‘acidolysis’. (A summary of this work in English is given in Pedro (1982)). Hydrolysis of minerals occurs through dilute solutions in the pH range 5–9.6 and may be total or partial. Total hydrolysis leads to the removal of all elements including silica, and to the precipitation of gibbsite and kaolinite minerals, whereas partial hydrolysis, under different conditions, leads to the formation of smectite minerals (Pedro, 1982). Acidolysis operates when the soil solution pH is <5 or has strongly complexing properties and can again be total or limited. Total acidolysis involves complete solubilization of minerals with no precipitation of Al. Limited acidolysis leads to the fixation of Al in octahedral and interlamellar positions in layer silicates. In general, acidolysis is associated with podzols, podzolic brown soils and acid brown soils of cold temperate climates, whereas hydrolysis is dominant in ferrallitic soils of the warm humid tropics and is prominent in warm temperate zones and in the dry subtropics (Table 1).

A similar synthesis relating to soil clays and weathering was described by Duchaufour (1960) who distinguished geochemical weathering under near neutral conditions with no organic acid anions and typical of tropical environments, and biochemical weathering under acid conditions with organic anions and typical of temperate climates. The former is characterized by neoformed clay minerals and the latter by clay minerals formed by transformation.

The overview of Jackson (1964) concerning the distribution, stability and weathering reactions of clay minerals in soils is, largely, consistent with the conclusions of the French soil scientists. Jackson described the dominant clay mineral types occurring in the Soil Orders of the taxonomy of the USDA Soil Survey Staff (1960). There was a predominance of micaceous, interstratified or
interlayered layer silicate minerals in soils such as Entisols, Inceptisols and Spodosols which occur widely in temperate climates. Clay minerals in soils of tropical climates such as Oxisols and Ultisols, were dominated by kaolinite and halloysite, in addition to gibbsite and sesquioxide. However, the occurrence of 2:1 to 2:2 intergrades was also indicated in both Oxisols and Ultisols and in the latter, vermiculite was also mentioned. In the weathering reactions described by Jackson the emphasis was placed very much upon transformation pathways (Fig. 1) and particularly on the role of hydroxy Al interlayers. Thus, it was envisaged that both smectite and vermiculite minerals could accumulate aluminium interlayers during extensive weathering, to the point where they approach a mineral resembling dioctahedral chlorite. A subsequent weathering step involving tetrahedra inversion could then give kaolinite (or halloysite), although kaolinite could also form from mica directly. Kaolinite formation was thus proposed as proceeding through 2:1→2:2→1:1 or through 2:1→1:1 pedogeochemical reactions. A major role for precursor crystals was also suggested in the growth of montmorillonite with hydroxy units of Al, Fe and Mg forming on pre-existing montmorillonite crystallites, which become silicatized to form new montmorillonite layers. Jackson (1964) concluded that “the almost universal occurrence of hydroxy cation units precipitated on the surfaces of layer silicates exposed to weathering is thus extremely fundamental to clay mineralogy and clay genesis”.

PRESENT PERSPECTIVES

This section is concerned with research bearing on the origin of clay minerals in soils, in a global context, which post-dates the seminal French and American work described above. In particular, it is of interest to consider whether the major conclusions reached need to be modified, refined or even abandoned in the light of the enormously increased volume of data now available on clay minerals in

![Fig. 1. Pathways for the formation of soil clay minerals as outlined by Jackson (1964).]
soils, often gained through the utilization of new techniques. Comprehensive reviews of the literature on soil clay minerals up until 1989 are available in the relevant chapters of the monograph entitled *Minerals in Soil Environments* (Dixon & Weed, 1989) published by the Soil Science Society of America. This forms a very useful starting point for the present review where the emphasis will be placed upon recent work.

**Mica minerals**

It is now clear, as it has been for some considerable time (Fanning *et al.*, 1989), that the micaceous minerals that occur in soil clay fractions have been inherited largely from parent rock or other material, where they originally formed under different P-T conditions from those pertaining at the earth’s surface. However, there is some evidence that mica may form pedogenically, but only in special circumstances. Thus, Niederbudde (1975) and Niederbudde & Kussmaul (1978) presented evidence to support the formation of mica-like clay minerals in fertilized loess-derived soils in Germany following K fixation by beidellitic smectites. The conversion of K-saturated smectites to mica-like products following extensive wetting and drying cycles in the laboratory, and involving change from turbostratic stacking to a semi-ordered structure (de la Calle & Suquet, 1988), suggests a possible mechanism for such a change. Similarly, it has been suggested that the clay micas found in the surface horizons of soils of arid environments, such as south west USA (Nettleton *et al.*, 1973) and Iran (Mahjoory, 1975) are pedogenic in origin. In both instances a transformation process was envisaged whereby it was possible to fix K in pre-existing smectites because of the hot and dry soil conditions.

The need for caution in proposing mica formation by pedogenesis is emphasized by previously mistaken interpretations concerning concentrations of clay mica in the surface horizons of Hawaiian soils. Initially, it was thought that this mica must be pedogenic because mica was undetectable in the sub-surface soil and the basaltic parent materials (Juang & Uehara, 1968) but later work using isotopic analysis showed conclusively that the mica mineral was older than the basalt flow from which the soil was derived and therefore could not be pedogenic (Dymond *et al.*, 1974). Aeolian deposition following tropospheric transport almost certainly explains the origin of the mica minerals in these Hawaiian soils. However, convincing evidence for neoformed mica in some Australian soils was presented by Norrish & Pickering (1983). A neoformational origin was indicated by the perfect platy hexagonal morphology of the clay mica and it is significant that the mineral is Fe-rich, in this respect resembling glauconite which forms in surface sediments under marine conditions. It is not clear, however, that the mica described by Norrish & Pickering (1983) is truly pedogenic in that the soil was derived from lacustrine sediments. It is possible that the mica may have formed in special conditions obtained during the drying out of the lake bed, rather than during soil formation on the lake bed sediments.

A general point made by Fanning *et al.* (1989) was that, at that time, there were very few detailed studies of the crystallochemistry of soil clay micas. This is true even now but there have been some interesting advances in the nineties brought about by the use of analytical electron microscopy, particularly by Robert and co-workers in Versailles. Although it may be agreed that virtually all soil micas are of inherited origin, it is still important to determine what changes may have occurred to the clay size micas under pedogenetic conditions, bearing in mind the importance attributed to micas as precursors of interstratified, expansible and interlayered minerals in soils. Robert *et al.* (1991) using HRTEM on impregnated sections of French soils derived from sediments, showed that many of the particles were exceedingly thin, recalling the fundamental illite particles described by Nadeau *et al.* (1984) in their studies of interstratified clays. Similar findings for clay-size micas in Iowa soils were arrived at independently by Laird & Nater (1993). Robert *et al.* (1991) went on to distinguish two types of mica mineral in the clay fractions of the soils they studied (Table 2). The first, described as micromica, occurs in

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Micromica</th>
<th>Illite (sensu stricto)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a b plane</td>
<td>micrometric</td>
<td>nanometric</td>
</tr>
<tr>
<td>Thickness</td>
<td>&gt;10 nm</td>
<td>~5 nm</td>
</tr>
<tr>
<td>K₂O</td>
<td>~10%</td>
<td>~7.5%</td>
</tr>
<tr>
<td>Charge</td>
<td>&gt;0.9</td>
<td>&lt;0.6</td>
</tr>
<tr>
<td>Structure</td>
<td>3-dimensional</td>
<td>3-dimensional</td>
</tr>
</tbody>
</table>
relatively thick particles of micrometric size with a full content of K and is highly charged. The second type is termed illite (sensu stricto) and occurs in thin particles of nanometric size, is somewhat depleted in K and is less highly charged. These two types of micaceous clays can exist in composite crystals, implying that there is a genetic relationship between them. Later work on the weathering of micaceous clays in soils (Romero et al., 1992; Aouidjit et al., 1996) does indeed show that particle microdivision is an important part of the process. Recent studies of Spanish soils developed upon highly micaceous parent materials presented chemical and structural evidence for the conversion of 2M1 muscovitic mica to the 1Md polytype with decreasing grain size (Martin Garcia et al., 1997).

The picture that is emerging from these investigations is that illitic clays may evolve in soil from the weathering of dioctahedral muscovitic mica minerals involving loss of K and layer charge and principally mediated by the physical process of particle microdivision.

Although the vast majority of micaceous soil clays are dioctahedral and aluminous, it should not be forgotten that trioctahedral micaceous soil clays do exist. Thus, Fordham (1990a) found that the weathering of granite in Australia, in conditions that led to the formation of lateritic soils, yielded trioctahedral micaceous clays in the initial stages of the weathering sequence. Again, Wilson et al. (1997) recorded trioctahedral micaceous clays accompanied by kaolinite in some alluvial soils from Nigeria. In both of these cases the trioctahedral clays appeared to be derived from biotite and this raises the question as to why vermiculitization of the biotite did not occur, as is commonly found in soils of humid-temperate climates. It may be relevant that in tropical soils, a direct biotite→kaolinite transformation can occur without an intervening vermiculite stage (De Kimpe & Tardy, 1968). It should also be noted that oxidized biotites retain their K more strongly than unoxidized biotites and are hence more resistant to vermiculitization. A possible explanation for the rather unexpected occurrence of trioctahedral micaceous clays in tropical soils is, therefore, that the precursor biotites were oxidized and weathered directly to kaolinite. Such kaolinite formation does not take place homogeneously throughout the biotite crystals, but tends to affect outer surfaces and cleavage planes exposed by exfoliation (Ojanuga, 1973) and in these circumstances it is conceivable that oxidized but otherwise relatively unaltered biotite fragments may eventually find their way into soil clay fractions.

**Vermiculite minerals**

The review of Douglas (1989) concerning the origin and formation of vermiculite minerals in soils emphasized a number of significant points. Firstly, there is general agreement that most vermiculites are considered to form by the weathering of mica as was illustrated early on by the work of Walker (1949). It is true that there is some evidence to support the formation of vermiculite by the weathering of chlorite, involving the preferential decomposition of the interlayer hydroxide sheet (Stephen, 1952; Ross & Kodama, 1976; Makumbi & Herbillon, 1972), or even through the decomposition of non layer-silicate minerals such as orthopyroxene (Basham, 1974), but in most instances there seems no need to look further than a genetic relationship with a precursor mica. This link is particularly obvious in the many studies that show vermiculite increasing towards the surface of soils with a concomitant decrease in the intensity of the 10 Å mica peak. The most detailed studies have focused on the biotite→vermiculite transformation.

The structure of vermiculite itself, following principally upon the fundamental studies of de la Calle and co-workers (references in de la Calle & Suquet, 1988), is well understood and the structures of both the silicate layers and the interlamellar space, as well as the various ways in which the layers may be stacked according to the saturating cation and relative humidity, have been elucidated. The vermiculitization of biotite involves a number of factors that have been explored in laboratory experiments. For example, the release of interlayer K to an external solution may take place by a diffusion process requiring that the concentration of K in the solution is below a critical level of 11 ppm (Newman, 1969). The vermiculitization reaction ceases if this critical level is exceeded. Numerous studies have shown that vermiculitization of biotite also involves the oxidation of structural Fe (Farmer et al., 1971; Vicente-Hernandez et al., 1983) and that in order to maintain overall electrical neutrality of the structure, this process results in the expulsion of Fe from octahedral sites. It is not clear whether oxidation of biotite can occur in soils without prior loss of K. Laboratory experiments indicate that loss
of K is necessary, but the natural occurrence of oxybiotites as well as evidence that a direct biotite to kaolinite weathering step is possible, as indicated in the above discussion on trioctahedral clay micas, suggest that these experiments may not necessarily apply to all soil conditions.

Whether or not previous vermiculitization of biotite is essential for oxidation of Fe to occur, the loss of Fe effectively leads to a more dioctahedral structure which in turn leads to a re-orientation of octahedral hydroxyl from a direction normal to the layer structure in unaltered biotite to a direction that is more inclined to the layer structure in oxidized biotite (Juo & White, 1969). This effectively places the interlayer K into a less negative environment so that it is consequently held more tightly within the structure (Norrish, 1973). The changes involving oxidation and loss of Fe have been suggested as explanations for the loss of layer charge in the biotite to vermiculite transformation, but other mechanisms may also be involved such as loss of hydroxyl ions.

Despite the weight of research effort on trioctahedral vermiculites, it is a fact that dioctahedral vermiculite is much more common in soil clays. This mineral, first discovered in British soils (Brown, 1953) occurs particularly in severely weathered soils such as Ultisols, Alfisols or even Oxisols and is also found commonly in the upper horizons of Spodosols (Douglas, 1989). Frequently, dioctahedral vermiculite is interlayered with non-exchangeable hydroxy AI. The reason dioctahedral vermiculitic minerals have not been so extensively studied is that they are found largely in clay fractions, where they are difficult to purify. A muscovite to dioctahedral vermiculite transformation analogous to the biotite to trioctahedral vermiculite transformation in the non-clay fractions of soils has rarely, if ever, been observed, and muscovitic mica cannot be readily vermiculitized in the laboratory without resort to extreme chemical techniques. Thus, there is no ready source of material with which to study experimentally such problems as the mechanisms involved in the loss of layer charge in dioctahedral vermiculite as there is for their trioctahedral counterparts.

Although the formation of dioctahedral vermiculite from a precursor dioctahedral mica seems probable in the majority of cases, it is of interest that other pathways have been suggested. Thus, in an early study, Barshad & Kishk (1969) presented evidence to support formation as precipitation products, but this proposal does not seem to have been widely accepted. Fordham (1990b) suggested, on the basis of analytical electron microscopy, that in a granitic weathering profile in Australia, trioctahedral vermiculite converted to dioctahedral vermiculite. It seems unlikely, however, that such a mechanism could explain the widespread occurrence of dioctahedral vermiculitic minerals in soils. Recent work by Aouidjit et al. (1996), however, using high resolution transmission electron microscopy (HRTEM) presented lattice images showing that vermiculitization of muscovitic-type mica did occur in very fine particles engendered by microdivision. In particular, they showed micrographs showing particles with a muscovitic core and a vermiculitic rim consisting of ~15 layers as well as small vermiculitic particles derived from muscovite consisting of 2 to 5 layers in thickness. This work suggests that very fine grained mica particles could be more susceptible to vermiculitization in the soil by reason of their higher specific surface and enhanced reactivity.

Smectite minerals

The factors that strongly influence the origin and formation of smectites in soils, as reviewed by Borchardt (1989), include low-lying topography, poor drainage and base-rich parent material, leading to favourable chemical conditions characterized by high pH, high silica activity and an abundance of basic cations. These conditions are met in many soils under temperate or cold climates or even in tropical climates where leaching is limited for various reasons, including low precipitation, a horizon in the soil profile that impedes the passage of water or a naturally high water table (Allen & Hajek, 1989).

Soil smectites are overwhelmingly dioctahedral, with only montmorillonite and beidellite being of any real importance. However, from a review of published chemical analyses, Wilson (1987) concluded that soil smectites were somewhat more Fe-rich than montmorillonite and beidellite sensu stricto and tended to show intermediate charge characteristics with many falling into the category of ferruginous beidellites. These observations imply that soil smectites could be somewhat different to those associated with bentonite and other geological deposits, which is perhaps not surprising considering the range of chemical conditions that may occur in soils as well as the three distinct pathways.
Origin and formation of clay minerals in soils

(inheritance, neoformation and transformation) that may be involved in soil smectite formation. Also, in the soil environment there is often no shortage of Fe, and pedogenic conditions often promote its solubilization and mobilization.

Wilson (1993) reviewed the literature in order to assess the relative importance of the three pathways for smectite formation in the Orders of Soil Taxonomy and some of the main points are as follows.

In Entisols, smectites are often inherited. Thus, inherited tri-octahedral smectite is found in the proto-soils forming from the 1980 Mount Saint Helens pyroclastic flow (Pevear et al., 1982; La Manna & Ugolini, 1987) and inherited smectite may occur in Entisols under an extreme range of climatic conditions as, for example, in the alluvial soils of the Blue Nile Clay Plains of Sudan (Wilson & Mitchell, 1979) and in Orthents and Psamment on altered basalt flows in the Faeroe Islands (Rutherford & Debenham, 1981). Neoformational smectite in Entisols may encompass a similar climatic range occurring in high pH footslope soils in southern India (Murali et al., 1978), in waterlogged horizons in typic Cryaquepts in the Bolivian Andes (Wilke & Zech, 1987) and in TorriFluvuents formed in wadi deposits under a xeric moisture regime in Saudi Arabia (Mashhady et al., 1980). Only one example was found where smectite in an Entisol was shown to form by transformation of a mica mineral, in this case celadonite in an altered basalt (Reid et al., 1988). This occurrence is likely to be exceptional.

Smectite may be the dominant clay mineral in Aridisols in the United States, and Dregne (1976) suggested that generally, environmental conditions are conducive to smectite formation. Smectite in arid soils has been noted in Iraq (Al Ravi et al., 1969), Iran (Abtahi 1977) and Saudi Arabia (Abahuseyn et al., 1980). Convincing evidence for a neoformational origin for smectite in saprolites weathering under arid conditions of the Negev desert of southern Israel was presented by Singer (1984). With regard to the smectitic argillic horizons in the Argid sub-order in the USA, most evidence suggests that these are Pleistocene relics related to a more moist climate and to eluvial deposition in the soil (Nettleton & Peterson, 1983). The smectite would, therefore, be regarded as being of inherited origin when viewed from the standpoint of current pedogenic conditions (Buol, 1965).

Smectites originating by transformation of layer silicates are unlikely to be of importance in Aridisols.

Smectites derived from inherited materials are common in Inceptisols, e.g. in the Fluventic Haplaquepts of the alluvial Indo-Gangetic plain (Islam & Lotse, 1986), in the Aquepts of the Bangladesh Holocene floodplain (Brammer & Brinkmann, 1977), in Typic Humaquept on glacial marine drift in north west Washington (Pevear et al., 1984) and in Japanese paddy soils derived from marine alluvium (Egashira & Ohtsubo, 1983). Neoformational smectites are also common in Inceptisols, a good example being that described from a Haplaquept in Nigeria by Bui & Wilding (1988) where in the prevailing aqueous environment, the silica translocated in the ferruginous soil has promoted the formation of an Fe-rich montmorillonite. Previous work by French scientists (Trauth et al., 1967; Paquet 1967; Tardy et al., 1974; Pedro et al., 1978) in this area of Africa also suggested a widespread neoformational origin for soil smectite. There seems to be little evidence in the literature for smectite to originate in Inceptisols by transformation processes.

Vertisols are dominated by smectite originating through inheritance or neoformation. Vertisols may have formed on alluvial plains where inherited smectite is of detrital origin such as in Sudan (Wilson & Mitchell, 1979), Turkey (Özkan & Ross, 1979; Güzeli & Wilson, 1981) and northern Uruguay (Rossignol, 1983). In these instances, the smectites are all Fe-rich with a high tetrahedral charge. Similar smectites are found in Vertisols developed upon basic igneous rocks, such as found in Kenya (Kantor & Schwertmann, 1974), Israel (Singer, 1971) and Jordan (Shadfan, 1983) where the clay mineral has a neoformational origin. Smectites originating by transformation of micas have not been conclusively demonstrated to occur in Vertisols although Badroui & Bloom (1990) and Graham & Southard (1983) have suggested the possibility. However, it would be surprising if a mica-to-smectite conversion is able to occur in the poorly drained conditions characteristic of Vertisols.

Smectites are common in Mollisols, particularly in the United States. An early study by Glenn et al. (1960) concluded that the smectite formed following weathering of micas and chloride and in a later review, Jackson (1965) concluded that smectites in many Mid-West soils had been formed by K depletion of di- and tri-octahedral micas or by the silication of polymerized sheets of
hydroxy-Al. In other words, the smectites are largely pedogenic, forming by a combination of transformation and neoformation. In a more recent study, Laird et al. (1988) investigated smectites in an Argialboll-Argiaquoll toposequence in Iowa to determine whether drainage differences influenced the chemistry and layer charge of the clay minerals. It was concluded that there was no drainage effect and that the smectite was likely to be a weathering product of inherited mica. It should be noted, however, that a comprehensive review of Canadian soils by Kodama (1979) reported that there had been only minimal transformation of the layer silicate clays in the Mollisols of the Prairie Provinces, with slight increases in expansible clays towards the surface in only a few instances. In Alfisols, the smectite of the argillic layer is surmised to have formed by different mechanisms. Where the parent material contains illite, it has been proposed that smectite forms by a transformation process, with vermiculite as an intermediate phase in some Ohio Hapludalfs (Smeck et al., 1968; Wilding et al., 1971). Generally, however, it has been proposed that smectite increases in the more poorly drained Aquafs located in low-lying areas (Allen & Fanning, 1983), presumably being of neoformational origin. Convincing evidence for the synthesis of smectite in a Lithic Hapludalf derived from serpentinite, which emphasized the role of poor drainage in the soil, was presented by Istok & Harward (1982). The albic horizon in Spodosols is the site of intense weathering and numerous studies have demonstrated that it is dominantly smectitic and related to the weathering of mica and sometimes chlorite. Ross & Mortland (1966) first showed that such smectite was beidellite, a finding since confirmed by others (Churchman, 1980; McDaniel et al., 1995). The aluminous nature of these clays, as well as the detailed study of Kodama & Brydon (1966) which showed that the smectite mineral in albic horizons is a three-component random interstratification of mica, vermiculite and smectite, represents convincing evidence that the smectite originates through transformation of the mica. A chlorite to smectite transformation is also possible as Spodosols with smectitic E horizons are sometimes found on parent materials which contain chlorite but no mica such as the tephras of the Pacific north west USA (Ugolini et al., 1991).

Smectites are usually found as a transitory phase in Ultisols particularly where there is an aeric moisture regime. Thus, Kantor & Schwertmann (1974) found that smectite increased with depth in some Kenyan Humults where it apparently formed as an initial weathering product of basic igneous rocks only to decompose in the surface horizons. Smectitic Ultisols have been more extensively studied in the USA where they may be developed upon smectite-rich shales, for example in the Alabama Coastal Plains (Karathanasis et al., 1986) and in Texas (Carson & Dixon, 1972). The above examples illustrate that Ultisol smectites are in the main inherited from suitable parent materials or may form, often ephemeral, through neoformation processes at the base of the profile. An origin through transformation is unlikely, although this has been suggested where Ultisols have formed on glauconitic sediments in the Mississippi Coastal Plain (Nash et al., 1988).

In view of the intensively weathered nature of Oxisols, it would be expected that smectites would be unlikely to occur. Certainly, this appears to be true for the diagnostic oxic horizon or for the plinthite found at depth. However, when lateritic profiles are studied in detail, smectite is sometimes found in the saprolite as an ephemeral phase (Nahon & Colin, 1982; Singh & Gilkes, 1993) or even in the so-called pallid zone (Singh & Gilkes, 1991). Considering the relative importance of the various modes of origin for smectite in Soil Taxonomy, the author’s conclusions are summarized in Table 3. Inheritance from smectitic parent material is judged to be probably of greatest importance where smectites are found in Entisols and Ultisols and of major importance in Aridisols, Inceptisols and Vertisols. A neoformation origin is of major importance in Inceptisols, Alfisols, Aridisols and Vertisols and of moderate importance in Entisols and Mollisols. Smectite formed by transformation is thought to be of major importance only in Spodosols and perhaps also in Mollisols and Alfisols although the evidence does not appear to be entirely conclusive. As a general comment, it seems to the author that too much importance may have been placed upon transformation mechanisms as a means of explaining the origin of soil smectites and not enough on neoformation and inheritance.

Chlorites and interlayered minerals

As indicated by Barnhisel & Bertsch (1989) there is no doubt that the occurrence in soils of primary chlorite of ferromagnesian, triocahedral nature, is
**Origin and formation of clay minerals in soils**

**Table 3. Origin of soil smectites in soil orders as assessed from the literature.**

<table>
<thead>
<tr>
<th>Soil orders</th>
<th>Inheritance</th>
<th>Neoformation</th>
<th>Transformation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entisols</td>
<td>+++</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>Aridisols</td>
<td>+++</td>
<td>+++</td>
<td>-</td>
</tr>
<tr>
<td>Inceptisols</td>
<td>+++</td>
<td>+++</td>
<td>++</td>
</tr>
<tr>
<td>Vertisols</td>
<td>+++</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>Mollisols</td>
<td>+</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>Alfisols</td>
<td>+</td>
<td>+++</td>
<td>++</td>
</tr>
<tr>
<td>Spodosols</td>
<td>-</td>
<td>+</td>
<td>+++</td>
</tr>
<tr>
<td>Ultisols</td>
<td>+++</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Oxisols</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

+++ Major importance  
++ Moderate importance  
+ Minor importance  
- No importance  

Due largely to inheritance. Primary chlorites are generally considered to be easily weatherable and Ross & Kodama (1976) provided convincing experimental evidence to show that chlorite could be changed into a regularly interstratified chlorite-vermiculite through chemical oxidation and selective extraction of the interlayer hydroxide sheet. Field occurrences of the vermiculitization of chlorite have been documented by Proust (1982), Proust *et al.* (1986) and by Buurman *et al.* (1988). Alternatively, chlorite may be completely decomposed without the formation of any crystalline product, as found by Bain (1977) in podzolic soils in Scotland.

The origin of so-called pedogenic chlorite and interlayered or inter-gradient minerals appears to be due entirely to transformation reactions involving the introduction of non-exchangeable hydroxy-Al polymers into the interlamellar space of pre-existing smectite, vermiculite or interstratified expansible minerals. The source of the interlayer Al could be direct involving acid attack of both tetrahedral and octahedral sheets of the layer silicate mineral, or indirect from the soil solution following weathering of aluminous minerals, like feldspars, or from decomposition of organic matter containing adsorbed Al. Both pathways, as well as the conditions under which they operate, have been confirmed by laboratory experiments. For example, it has long been known that acidified smectite spontaneously converts to an Al-saturated form whereby H⁺ is replaced by structural Al. Again, the decomposition of trioctahedral micas by mineral acids and non-complexing organic acids was shown by Vicente *et al.* (1977) to result in the formation of Al hydroxy interlayered vermiculite, whereas the effect of highly complexing organic acids often resulted in complete breakdown of the mineral structure. Many experiments have shown that water-soluble polynuclear Al complexes are strongly adsorbed by expansible minerals to yield hydroxy-interlayered clays (e.g. de Villiers & Jackson, 1967), and that the original exchangeable cations are quantitatively displaced (Brown & Newman, 1973). According to Vicente *et al.* (1977) Al hydroxy-interlayered clays in soils mainly relate directly to the decomposition of the layer silicate mineral itself. They further concluded that interlayered minerals are the most common clay minerals in the soils of Northern Europe and North America and that they may result not only from natural soil processes but may also be promoted by the application of fertilizers and the deposition of acidity from the atmosphere.

Extensive consideration has been given to these minerals in the latter context, particularly because the stability of the interlayers is pH dependent, being favoured in the pH range extending approximately from 4.0 to 5.8 (Barnhisel & Bertsch, 1989). Hydroxy interlayered clays can thus serve as sources and sinks of mobile Al depending upon soil pH conditions. Other soil factors that promote their formation include low organic matter content, oxidizing conditions and frequent wetting and drying cycles. These conditions are met in a wide range of soils, and interlayered minerals are thus commonly encountered. They may be most typical of highly
weathered soils and they certainly occur in Ultisols (Karathanasis, 1988) and even in Oxisols.

However, some emphasis has recently been placed on the occurrence of hydroxy interlayered clays in the podzolic soils of catchments on slowly weathering parent rocks and where surface waters are considered to be at risk from acidification. For example, April et al. (1986) described hydroxy Al interlayered vermiculite, primarily in B horizons of soils from 26 Adirondack catchments. Olson (1988) suggested that such minerals play a key role in neutralizing potential increases in acid deposition in catchments in montane areas of Maryland and Virginia. In Scotland, Bain et al. (1990) presented evidence to show the occurrence of these minerals in the podzolic soils of three upland catchments and also demonstrated the pH dependency of the interlayers. The extent of the interlayering declined with falling pH and complete removal was detected below pH 4.3 in the E horizons.

Whether acid precipitation has a major role in bringing about such changes is rather difficult to determine, although from laboratory experiments it is known that changes in hydroxy interlayering of expansible phases can occur very rapidly. In most soil studies there is no historical time scale to assess the rate of change and to judge whether changes coincide with marked periods of acid deposition.

Hydroxy Al interlayers in expansible clay minerals may become silicated as a result of interactions either outside or within the interlamellar space (Lou & Huang, 1993) leading to the formation of an intergradient vermiculite-kaolin mineral (Wada & Kakuto, 1983) or to kaolinite directly (Karathanasis & Hajek, 1983) as will be discussed below. Interlayered minerals are mainly dioctahedral, representing the nature of the precursor expansible minerals, but Al interlayered trioctahedral vermiculites may also occur during biotite weathering (Kato, 1965; Wilson, 1965, 1966).

Interstratified minerals

These minerals primarily represent intermediate transformation products, mainly involving mica, chlorite and an expansible phase, either smectite or vermiculite (Sawhney, 1989), although there are many examples of interstratified minerals in soils originating by inheritance. Both regularly and randomly interstratified clays either of dioctahedral or trioctahedral nature, may occur in a wide variety of soil types. However, because they represent intermediate transformation products, they would perhaps not be expected to occur widely in old highly weathered soils.

Wilson & Nadeau (1985) reviewed the types of interstratified minerals that are characteristic of weathering environments. Hydrobiotite, consisting of large crystals of regularly alternating biotite and vermiculite layers and derived largely from the weathering of biotite, represents a classic case of an interstratified mineral that was observed early on. It has been found many times in soils (e.g. Coleman et al., 1963; Wilson, 1970; Kapoor, 1972), appears to form relatively rapidly (Wilson & Nadeau, 1985), and can be synthesized in the laboratory using dilute solutions to exchange interlayer K (Rausel Colom et al., 1965) or by oxidation of structural Fe during cation exchange (Farmer & Wilson, 1970). Norrish (1973) proposed a convincing mechanism for the formation of hydrobiotite involving sympathetic re-orientation of the hydroxyls on either side of the octahedral sheet. Removal of K from one side of the silicate layer increases the angle of the (OH) bond direction on that side of the octahedral sheet, but causes a decrease in the angle of the (OH) band on the other side. This results in the interlayer K on this side being placed in a more negatively charged and hence more stable environment and alternating interlayer regions are built up within single crystals where K is either depleted or held very strongly. Regularly interstratified dioctahedral mica-vermiculite is also found in soil clay fractions particularly in the upper horizons of podzolic soils (Churchman, 1978, 1980; Wilson et al., 1984) and presumably formed from dioctahedral mica by a mechanism similar to that of hydrobiotite formation. However, the dioctahedral mineral does not seem to occur in non-clay fractions. Further weathering of vermiculitized mica results in a lower negative charge on the silicate sheet leading to the formation of smectitic products of both a trioctahedral (Isma'il, 1969; MacEwan, 1954) and dioctahedral nature (Churchman, 1980). In the latter instance a regularly interstratified mineral may form with the XRD characteristics of rectorite.

Regularly interstratified chloritic material may occur in soils as a consequence of the vermiculitization of chlorite, although the mechanism involved is not clear. A regularly interstratified chlorite-vermiculite yielding a high spacing of 28 Å with many lower orders was described by Johnson.
(1964) from the C horizon of a soil developed on metamorphosed basalt in Pennsylvania and a similar mineral was reported by Herbillon & Makumbi (1975) in a recent tropical soil derived from chlorite-schist in Zaïre. In the latter instance, at least, it was evident that the chlorite from which the interstratified product developed was already partly vermiculitized (Makumbi & Herbillon, 1972) but the experiments of Ross & Kodama (1976) demonstrate that some true chlorites will break down to a regularly interstratified product.

The formation of regularly interstratified minerals and thence of vermiculite or smectite must involve transition through randomly interstratified phases but it is of interest that these phases appear to be relatively transitory and do not seem to persist in soil profiles. Comparison with diagenetic sequences is striking where great thicknesses of argillaceous sediment may be characterized by randomly interstratified illite-smectite which converts slowly with depth and increasing temperature and pressure to more ordered forms including rectorite-like structures (R1) and more long-range ordered structures (R2 and R3). In the author’s view, there is little evidence that the long-range ordered illite-smectites nor indeed the precursor randomly interstratified illite-smectite clay form as stable phases in soils. This is perhaps not surprising in view of the increasing P–T conditions required for this transformation in diagenesis. One of the few papers to show that interstratified illite-smectite will form at earth surface conditions is that of Bergkraut et al. (1994) in a study of weathering in basic pyroclastics. However, the illite-smectite differs from that found in diagenetic sequences in that it is Fe-rich, recalling the illitic clays formed at earth surface conditions.

Interstratified kaolin-smectite is now becoming more widely reported in soils. It appears to be particularly common in soils derived from basic igneous rocks and, for example, was identified by Norrish & Pickering (1983) as the only clay mineral in ~40 Australian soils developed upon such parent material. The occurrence of kaolinite-smectite in a red-black toposequence in Burundi (Herbillon et al., 1981) suggests a genetic link between the black smectitic soils in the poorly drained footslopes and the red kaolinitic soils on the freely-drained hilltops. More recent work shows that the kaolin mineral in this kind of interstratification may be halloysite (Delvaux et al., 1990) and that the expansible mineral can be vermiculitic (Bühmann & Grubb, 1991). The transformation of smectite to interstratified halloysite-smectite was shown in weathered acid clay in Japan to involve the crystallization of excess silica in a separate phase (Watanabe et al., 1992).

Although Wilson (1987) concluded that the concept of fundamental particles was probably not of general applicability in the interpretation of interstratified smectitic clays in soils, Robert et al. (1991) showed by HRTEM that these particles did exist in soils and, from these observations, Aouidjit et al. (1996) distinguished two types of interstratification described as ‘structural’ and ‘textural’. The former involves relatively large crystals of micrometric lateral extent, variable thickness and three-dimensional structure. Textural interstratification consists of associations of elementary particles with a random two-dimensional structure and is of variable thickness and lateral extent.

**Kaolin minerals**

Dixon (1989) reviewed the origin and formation of kaolin minerals and showed that precipitation from solution of kaolinite required acid conditions with moderate silica activity and small amounts of base cations. Kaolinite was synthesized from hydroxy Al interlayered montmorillonite by Poncelet & Brindley (1967) but only under hydrothermal conditions. Neoformation and transformation mechanisms are, therefore, feasible. Kaolinite is generally found as a minor constituent in young soils, such as those derived from glacial material, and by implication is considered to require >10,000 years to form. It may be an abundant constituent of soils developed upon old geomorphic surfaces.

Compared with the kaolinites of geological deposits, soil kaolinites are often of a smaller particle size, tend to be highly disordered and may contain Fe as an isomorphous substituent. They may also be interstratified with smectites. The pioneering work of Angel et al. (1974), Jones et al. (1974) and Meads & Malden (1975) using EPR spectroscopy showed that Fe$^{3+}$ could be substituted for Al in the octahedral sheet of kaolinite and that there were different sites of substitution. For the kaolinite of tropical soils Herbillon et al. (1976) demonstrated a relationship between Fe content, crystallinity index and EPR spectral features, particularly the intensity and asymmetry of the $g_{eff}$~4 band. Mendelovici et al. (1979) and
Cantinolle et al. (1984) showed convincingly by IR spectroscopy that structural Fe occurs in kaolinite in lateritic and bauxitic soils. It seems, therefore, that soil kaolinites (like soil smectites), may be different from the kaolinites of geological deposits, presumably because of the large amounts of Fe in the pedogenic environment combined with conditions that promote the mobilization of this element, at least on a local scale. The characterization of kaolinite by EPR was developed further by Müller & Calas (1993) who showed that the technique could be used to distinguish different environments of formation according to total Fe$^{3+}$ content, distribution among different sites (Fe(i) and Fe(ii)), concentration and types of paramagnetic defect centres (PDC) and presence or absence of sorbed species. Soil kaolinites yielded a high intensity Fe$^{3+}$ EPR signal, showed an interdependence between Fe(i) concentrations and crystalline order, low intensity signals from PDCs and lack of resonance due to sorbed species.

Halloysite occurs primarily in youthful volcanic-derived soils, but it also forms from primary minerals in tropical soils or pre-glacially weathered materials. Most commonly the mineral is of a tubular or spheroidal morphology. Platy halloysite is rare. Although similar to kaolinite in its layer structure, it is important to realize that halloysite is fundamentally different in that it has a 2-layer structure whereas kaolinite has a 1-layer structure (Bailey, 1989). This means that a halloysite to kaolinite conversion cannot take place by a simple structural transformation but would require a recrystallization process, as was found experimentally by La Iglesia & Galán (1975). The reasons for the formation of spheroidal halloysite as opposed to tubular halloysite are not clear, although it may be that high levels of silica supersaturation are required (Tomura et al., 1985) and the appropriate physical conditions provided by pumice grains. Tubular halloysite may eventually form from the spheroidal form (Sudo & Yotsumoto, 1977). It is worth noting too that, like kaolinite, halloysite may contain structural Fe (Wada & Mizota, 1982), particularly where it has formed from the weathering of ferruginous material (Dao Cho & Mermut, 1992). There is evidence gained by X-ray powder electron spectroscopy that spheroidal halloysite may have a greater Fe content than the tubular form (Bailey, 1989).

Although most studies would seem to indicate that kaolinite in soils has formed by neoformation, there is evidence that a transformation mechanism is feasible. Thus, Wada & Kakuto (1983) presented evidence to show that in Korean Ultisols, kaolinite had formed via an intergradient vermiculite-kaolin mineral involving tetrahedral inversion and attachment to the hydroxy Al interlayer sheet. A similar mechanism for the formation of kaolinite from hydroxy interlayered smectite was proposed by Karathanasis & Hajek (1983).

**FUTURE PERSPECTIVES**

This overview has highlighted the necessity for further research to address many specific questions and gaps in our knowledge as to the origin and formation of clay minerals in soils, as well as further characterization of the detailed nature of the minerals themselves.

With regard to soil micas, it is clear that further detailed studies are needed on their crystallochemistry and on the changes that occur with decreasing particle size, particularly the possible conversion from the 2Md to the 1Md polytype and the role of microdivision in bringing about fundamental particles, or at least particles that are more susceptible to vermiculitization. At this stage the relative roles of increased basal surface as opposed to edge surface is unclear. The question of formation of pedogenic micas is one that remains to be resolved. Even the reality of the process seems uncertain, a point that could be clarified by K/Ar dating and other studies of carefully fractionated samples.

For reasons that have been given above, most of the information that we have on vermiculite refers to trioctahedral vermiculite, despite the fact that dioctahedral vermiculite is much more common in soils. Therefore, we need to know more about the mechanisms of formation of dioctahedral vermiculites, particularly the changes in chemistry involved in the reduction of layer charge during evolution from the parent micas. In this instance, oxidation of structural Fe is unlikely to play as significant a role as it does in trioctahedral vermiculites. There is very little information on the stacking sequence of the silicate layers and the organization of the interlamellar space for dioctahedral vermiculites, in comparison with that available for their trioctahedral counterparts.

Although there is good evidence that soil smectites may form through both transformation and neoformation pathways, more information is
needed with regard to possible differences in chemistry and structure that these two modes of origin may engender. For example, transformation smectites should, in principle, be beidellitic, as they should inherit Al-substituted tetrahedral sheets from their parent micas, and they might also be expected to show semi-random stacking (again an inherited feature from their precursor vermiculites), as opposed to turbostratic stacking. The mechanism involved in the reduction of layer charge, and the distinction from the vermiculite from which the transformation smectites are derived is another area for further research. With regard to neoformation smectites, confirmation is required by direct methods that many can indeed be described as ferruginous beidellites, as is suggested by chemical analyses. However, such analyses can be misleading because non-structural Fe oxides may be resistant to deferration techniques. There may perhaps be scope to use EPR techniques for the assessment of environmental conditions for smectite formation, in the same way as they have been used in kaolinite studies. In any event, it would seem that detailed comparison between transformation and neoformation soil smectites would be worthwhile.

In the author’s view, there is ample scope for further research into the interstratified clay minerals that are truly characteristic of the pedogenic environment. It seems that the present picture is somewhat confused by the implied formation in soils of the same types of interstratified illite-smectites that are so common in the diagenetic environment. This is unreasonable, bearing in mind that the latter are characterized by increasing conditions of temperature and pressure and that there is a dearth of evidence to show that diagenetic-type illite-smectites really do form in the soil. It is true that K saturation combined with wetting and drying cycles can convert smectite to an illite-like product in the laboratory, but this product is not illite sensu stricto and K saturation is unlikely to occur in the soil except under arid conditions. Further research is required to address this issue. Other specific problems that might be addressed include the mechanisms for the formation of regularly interstratified dioctahedral mica-vermiculite and for interstratified kaolinite-smectite and halloysite-smectite in soils. Further information is required too on the possible role of microdivision and the creation of fundamental particles in promoting interstratification as suggested by Robert and his colleagues.

For the kaolin minerals, it seems that genesis through a transformation mechanism from hydroxy-interlayered minerals involving tetrahedral inversion of silicate layers requires further documentation and assessment as to its general applicability in soils. Also, there is no question that a better understanding is required of the conditions of formation of halloysite as opposed to kaolinite, in soils, and in particular the formation of spheroidal halloysite vs. tubular halloysite. In these endeavours, it seems certain that the use of advanced spectroscopic techniques such as EPR will play an ever increasing role.

With regard to the global overviews of clay mineral formation in soils described at the beginning of this paper, it still seems that they can be regarded as generally valid and that they provide a useful frame of reference. It may be, however, that uncertainty still remains about the relative importance of transformation and neoformation in both general and specific instances. Figure 2 summarizes the author’s views on this matter in terms of primary and secondary pathways for clay mineral formation in soils in a global context. It can be seen that soil smectites and kaolinites are regarded primarily as the products of neoformation. For smectite, an important, but secondary pathway, applicable mainly to Spodosols and perhaps to Mollisols and Alfisols, is through the weathering of inherited mica (and to a lesser extent chlorite) and may or may not involve the intervention of hydroxy-Al interlayered minerals. Kaolinite too, although primarily a soil mineral arising from neoformation, also forms through other pathways of transformation, namely direct weathering of mica, conversion of hydroxy-Al interlayered vermiculite or smectite and weathering of smectite through an intermediate kaolinite-smectite interstratified phase. Halloysite is regarded primarily as a product of neoformation, although it seems that it may also form through an interstratified halloysite-smectite as a result of smectite weathering.

Finally, it is worth emphasizing that the question of the origin and formation of clay minerals in soils is not an exercise of mere academic interest but may relate directly to soil properties and behaviour in numerous ways. A better understanding of the detailed nature of clay minerals in soil and the way in which they relate to overall soil chemical and physical properties, combined with a knowledge of their mode of formation, provides a sound basis for
Fig. 2. Pathways for the formation of clay minerals in soils. Note that the major pathways for smectite and kaolinite are considered to be through neoformation as represented by thick arrows. Transformation pathways to smectite occur via interstratified and interlayered phases. A direct conversion of mica to kaolinite can also occur but this may not be regarded as a true transformation because of the structural differences between the minerals, hence the dotted line.

ACKNOWLEDGMENTS

I am indebted to the Clay Minerals Group Committee for the invitation to prepare this review paper. This work was supported by the Scottish Office Agriculture Environment and Fisheries Department.

REFERENCES

Origin and formation of clay minerals in soils

Wisconsin.


Soil Survey Staff (1960) *Soil Classification: A Comprehensive System*. USDA.


