Table 1 5′-terminal nucleotide analysis of *in vitro* labelled mitochondrial poly(A)-containing RNA 16 and 12S rRNA

RNA species	Percentage of total ^{32}P radioactivity* AMP	CMP	GMP	UMP
12S rRNA	91.8	1.9	2.0	4.2
Poly(A)-containing RNA 16	83.6	1.3	8.2	7.0

The 5′-terminal nucleotide was determined by exhaustive digestion of the *in vitro* labelled RNAs with nuclease P1 (ref. 14; Calbiochem) and fractionation of the products on PEI-cellulose TLC plates, as previously described[11].

* The values refer to the radioactivity which migrated from the origin. The percentage of the input radioactivity recovered from the origin was, respectively, 1.0 and 1.6% for 12S rRNA and RNA 16.

nucleotides in the RNA 16 preparation, during the 5′-end labelling reaction, or may be due to the formation, during the enzymatic or chemical degradation of the RNA, of oligonucleotides with different phosphate end-groups (2′,3′-cyclic phosphate, 2′ and 3′ phosphate). Similar bands migrating in abnormal position have been observed, previously[11] and in the present work, in the sequencing gel patterns for 12S rRNA. The reading of the first few nucleotides of the sequence near the 5′-end was made somewhat difficult by the presence of these extra bands; beyond the fifth nucleotide, however, the sequence could be read unambiguously. The sequence obtained is shown in Fig. 4, aligned with the DNA sequence of the COII gene region. The entire 29-nucleotide stretch of RNA 16 which has been sequenced appears to be colinear with the DNA sequence. The striking result is that the 5′-end of the RNA corresponds precisely to the first nucleotide of the COII coding sequence.

The observation that the COII mRNA starts directly at the initiator codon raises interesting questions about the mechanism by which the mitochondrial ribosomes attach to this messenger. In all eukaryotic mRNAs analysed so far there is a stretch of variable length which precedes the initiator codon and which is assumed to be involved in ribosome attachment[2,3]. Similarly, in *Escherichia coli* mRNAs there is a 5′-noncoding segment containing a ribosome binding site[17,18]. A single exception to this rule has been described, namely that of the λ repressor mRNA involved in repressor maintenance[19]. In this case, it was argued that the lack of a strong ribosome binding site could account for the low efficiency of translation of this mRNA. In the case of the mitochondria from human cells (and probably from other animal cells) it is reasonable to think that the special features of their ribosomes make them suitable for binding directly to the initiator codon. It will be interesting to see whether the lack of a 5′-noncoding stretch is a general feature of mitochondrial mRNAs in HeLa cells. Preliminary observations indicate that another mitochondrial poly(A)-containing RNA in these cells, RNA 15, starts with an AUG, which may be the initiator codon for the polypeptide coded by this RNA (unpublished observations).

This work was supported by NIH grant GM-11726. We thank S. Crews for valuable discussions and A. Drew for technical assistance.

Received 23 April; accepted 26 June 1980.

1. Barrell, B. G., Bankier, A. T. & Drouin, J. *Nature* **282,** 189–194 (1979).
2. Hagenbüchle, O., Santer, M., Steitz, J. A. & Mans, R. J. *Cell* **13,** 551–563 (1978).
3. Kozak, M. *Cell* **15,** 1109–1123 (1978).
4. Ojala, D. & Attardi, G. *Plasmid* **1,** 78–105 (1977).
5. Southern, E. M. *J. molec. Biol.* **98,** 503–517 (1975).
6. Alwine, J. C., Kemp, D. J. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5350–5354 (1977).
7. Amalric, F., Merkel, C., Gelfand, R. & Attardi, G. *J. molec. Biol.* **118,** 1–25 (1978).
8. Berk, A. J. & Sharp, P. A. *Cell* **12,** 721–732 (1977).
9. Berk, A. J. & Sharp, P. A. *Proc. natn. Acad. Sci. U.S.A.* **75,** 1274–1278 (1978).
10. Ojala, D. K., Merkel, C., Gelfand, R. & Attardi, G. (in preparation).
11. Crews, S. & Attardi, G. *Cell* **19,** 775–784 (1980).
12. Attardi, G. *et al. ICN-UCLA Symp. molec. cell. Biol.* **15,** 443–469 (1979).
13. Gelfand, R. thesis, California Institute of Technology (1980).
14. Fujimoto, M., Kuninaka, A. & Yoshino, H. *Agr. Biol. Chem.* **38,** 1555–1561.
15. Donis-Keller, H., Maxam, A. M. & Gilbert, W. *Nucleic Acids Res.* **4,** 2527–2538 (1977).
16. Simoncsits, A., Brownlee, G. G., Brown, R. S., Rubin, J. R. & Guilley, H. *Nature* **269,** 833–836 (1977).
17. Shine, J. & Dalgarno, L. *Biochem. J.* **141,** 609–615 (1974).
18. Steitz, J. A. & Jakes, K. *Proc. natn. Acad. Sci. U.S.A.* **72,** 4734–4738 (1975).
19. Ptashne, M. *et al. Science* **194,** 156–161 (1976).

Structural and functional diversity in 4-α-helical proteins

Patricia C. Weber & F. R. Salemme

Department of Biochemistry, New Chemistry Building, University of Arizona, Tucson, Arizona 85721

Protein crystallographic studies show that many structural arrangements appear as common features among proteins which are otherwise unrelated in sequence or function. One of the more recently recognized recurring protein structural motifs is a nearly parallel arrangement of four α-helices to form a sequentially connected left-twisted bundle[1]. We describe here the geometrical properties of these structures and suggest how they relate to the functional and aggregate properties of these molecules.

Currently known proteins which incorporate a 4-α-helical bundle as their dominant structural feature include (1) the haemerythrin subunit and related monomer myohaemerythrin, which are non-haem iron-containing oxygen transport proteins derived from marine worms[2–4], (2) the apoferritin monomer, which associates to form a hollow, 24-subunit cube-octahedral aggregate functioning in iron storage[5], (3) tobacco mosaic virus (TMV) coat protein[6,7], (4) the monomeric haem protein, *Escherichia coli* cytochrome b_{562} (ref. 8), and (5) the dimeric cytochrome *c′* from *Rhodospirillum molischianum*[9].

Figure 1 schematically compares the known 4-α-helical subunit structures in cylindrical projection. Although the molecules clearly differ in their α-helix, connecting loop, and overall chain lengths, they share a similar overall folding plan in that each is composed of four sequentially joined α-helices (A–D) arranged in a roughly fourfold symmetrical bundle having a net left-handed twist. As shown in recent treatises on helix packing[10,11], individual α-helices can be described, to a first approximation, as twisted parallelepipeds having a square cross-section. It is consequently obvious that the observed arrangement of four nearly-parallel α-helices to form an array having a square cross-section minimizes the accessible surface area of the structure as a whole[12]. Additionally, if it is assumed that these molecules fold from an intermediate state having preformed α-helices[10,13], it can be seen from inspection of Fig. 1 that all sequential helix pairs and intervening connecting loops have a net right-handed superhelical character. This is a well known chiral effect observed in several other protein structural domains[14–23] and is presumed to arise through some intrinsic property of extended polypeptide chains that confers stability on conformations having overall right-handed supercoil characteristics[24,25]. In the present case, these right-handed supercoiled connections provide the most efficient connectivity pattern for the bundle, since alternative left-supercoiled connections would necessitate the formation of a much larger loop passing around the entire molecule (Fig. 1).

Analysis of the available coordinate[26] or literature data[27] shows the helix packing in these molecules to be quite intimate, the mean adjacent interhelix axis distance of closest approach being 9.6 Å, with a standard deviation of 1.4 Å. Interhelix angles are remarkably similar, the mean interhelical angles for adjacent and diagonally related helix pairs being $162.0 \pm 5.8°$ and $-158.2 \pm 7.0°$ (mean ± s.d.), respectively[10]. The near constancy of the adjacent pair interhelix axis angle at ~18° (180° − 162°) reflects the fact that the individual helices can only make intimate and extended packing arrangements when their axes are relatively tilted at an angle equal to twice the α-helix pitch angle (~9.0°; ref. 28). Although a detailed description of all the 4-fold symmetric, sterically acceptable 4-α-helical arrangements lies outside the scope of this communication, it can readily be demonstrated that in fact there exists a variety of

0028–0836/80/360082—03$01.00

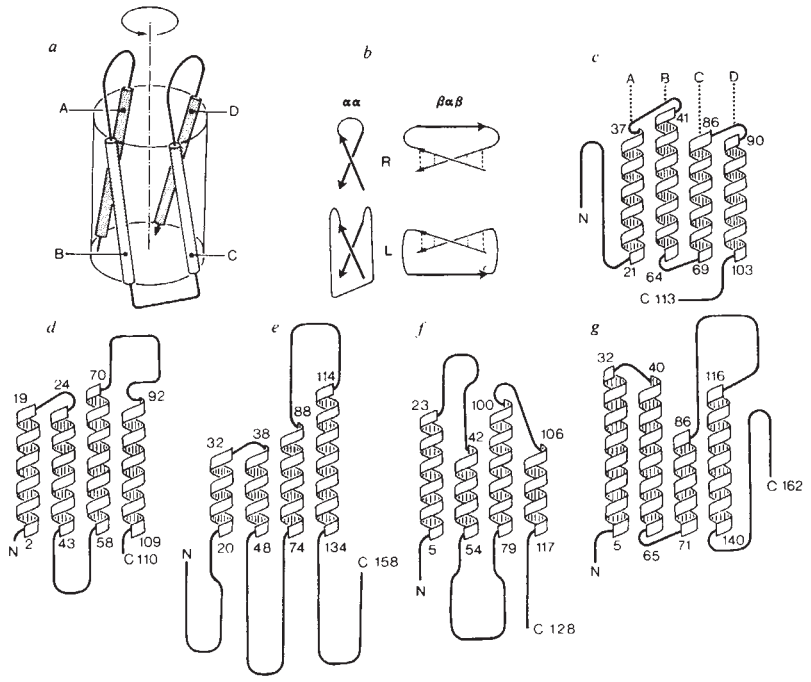

Fig. 1 Schematic representations of the 4-α-helical protein structures. *a*, Schematic representation of helical packing; *b*, schematic representation of chiral connectivity; *c–g*, examples of proteins showing these characteristics: *c*, haemerythrin and myohaemerythrin; *d*, cytochrome b_{562}, *e*, TMV protein; *f*, cytochrome *c'*; *g*, apoferritin. Note that the connectivity pattern shown here and in Fig. 3 of the apoferritin monomer is one of two possibilities (the alternative being left-handed[5]) either of which appears equally consistent with the current electron density map interpretation (P. M. Harrison, personal communication). Each molecule is illustrated as a cylindrical projection of a common arrangement of four α-helices which are sequentially connected to form a bundle having an overall left-handed twist about the approximate 4-fold bundle axis of symmetry. This arrangement results in a right-handed (R) supercoiled connectivity pattern between adjacent helices, a feature previously observed in several other protein structural domains[14–23]. In *b*, for example, we compare right- and left-hand connected helix pairs with analogous β-α-β domains.

related packing arrangements which are both consistent with the observed geometry and have overall symmetric interactions. Further, as can be shown by an appropriate helix-net representation[28], the interhelix angle of 18° is a value which is uniquely associated with the attainment of regular periodic interactions between all helix pairs in a 4-α-helical bundle. In Fig. 2*a*, we show a packing model[28] of a representative symmetric arrangement which illustrates these points, and a similar representation of the actual packing interactions in cytochrome b_{562} is shown in Fig. 2*b*. These structures are clearly related, and so the previously reported occurrences of high internal symmetry in 4-α-helical structures[27,29] would seem to reflect the symmetrical packing requirements for helices in these bundles.

Because the α-helices in the 4-α-helical bundle structures are straight, they must necessarily diverge from their point of closest interhelix axis approach. Figure 2*c–f* illustrates this behaviour for several 4-α-helical proteins, and demonstrates both that the helical bundles of these structures have been truncated near their points of closest interhelix axis approach, and that the spatial divergence of the helices, which results as a consequence of the helix packing requirements, serves to create the internal binding pocket for the incorporated prosthetic groups.

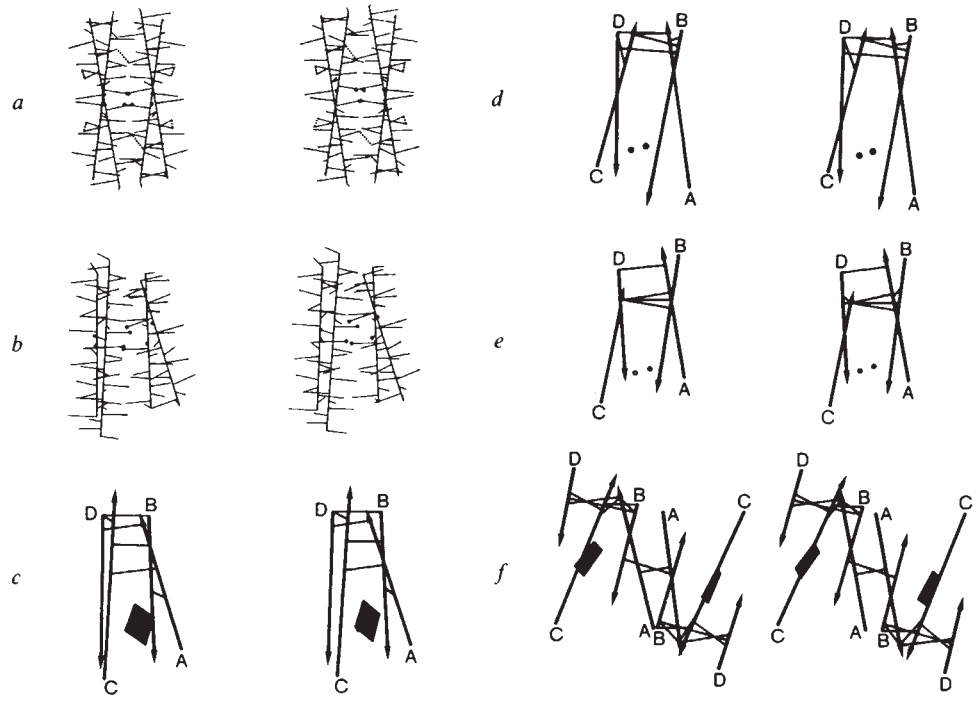

Fig. 2 Packing geometry in 4-α-helical proteins. *a*, A 'broomstick' packing model[28] in which each helix is represented as a helix axis line with 5.6 Å radial vectors[10] passing through Cβ. All adjacent helix pairs have interhelix axis angles of 18° and interaxis close-approach distances of 9.6 Å, so that the structure as a whole has the average geometrical properties of the observed structures. The structure shown has nearly equivalent local packing interactions (for example, dotted connection) which approximately repeat in the direction of the long axis of the bundle. (We note here that the interactions between the residues of straight α-helices can, at best, be only quasi-equivalent along the bundle axis. This situation differs from that observed in superhelical fibrous structures[28].) The arrangement as a whole has 222 point group symmetry, as a consequence of which each face of the bundle presents a packing surface with a diad axis of symmetry. The structure shown is only one of many possible symmetrically related arrangements which share a common overall geometry. *b* Shows a broomstick representation of the helix packing interactions in cytochrome b_{562} (ref. 8). The packing interactions in this molecule are not as regular as the model in *a*, because the residues of the real protein have different packing volumes, and the helical bundle comprising this molecule has been truncated near the region of closest interhelix approach (see *c*), with a concomitant reduction in apparent packing symmetry. Nevertheless, comparison of *a* and *b* clearly shows the helix packing in these structures to be related. *c–f* Show helix axis representations (including close approach interhelix vectors) for several molecules. As shown, all are truncated versions of a generalized 4-α-helical bundle model, which incorporate prosthetic groups in the internal cavity created by the spatial divergence of the individual helices away from their points of closest approach. An arrow indicates the N to C polypeptide chain sense in each helix. *c*, Cytochrome b_{562} (ref. 8). The position of the noncovalently bound haem group is indicated by a solid square. *d*, Myohaemerythrin[2]. Solid circles indicate the location of the bound iron atoms. *e*, Haemerythrin[4]. Iron positions are largely conjectural. *f*, Cytochrome *c'* dimer[9] showing positions of the covalently bound haem prosthetic groups.

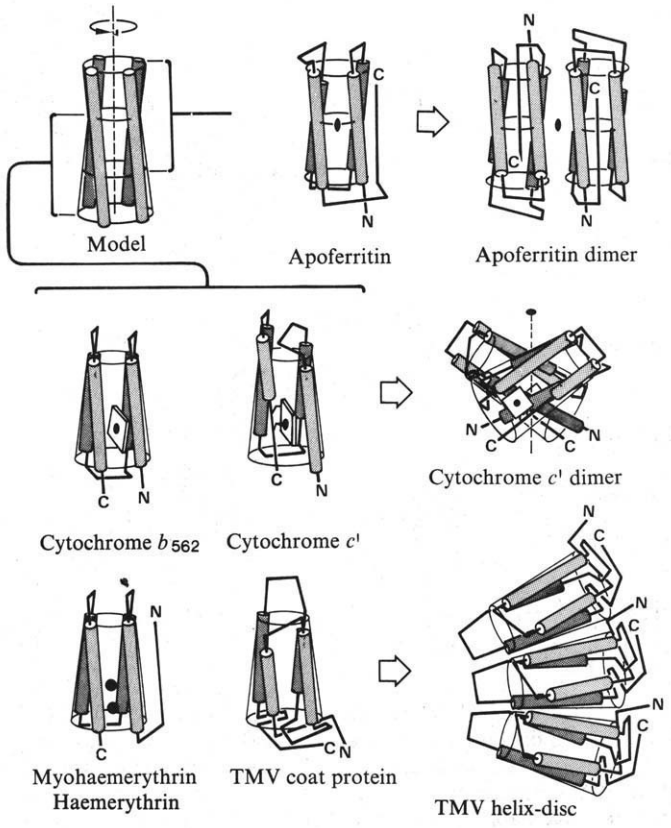

Fig. 3 Schematic representations of the 4-α-helical proteins illustrated as sub-domain arrays of a generalized 4-α-helical divergent bundle model. Broad arrows indicate the relationship between the monomeric and aggregate structural properties of the respective molecules. Details are given in the text.

Figure 3 shows the known proteins of this class illustrated as truncated variations of a generalized 4-α-helical packing model, and summarizes some additional features of their structural and aggregate properties. For example, in addition to the molecules described in Fig. 2, TMV coat protein also clearly appears to be a truncated divergent bundle structure, which instead of incorporating a prosthetic group, has a β-sheet at its divergent end which serves to 'brace' the helices apart[6]. The apoferritin monomer, in contrast, incorporates neither of these features[5] and so its helices pack more symmetrically with an equal extent of interhelical divergence at either end of the bundle. As a consequence, the most extensive pairwise monomer interaction in the cube-octahedral aggregate structure is a symmetrical dimer (Fig. 3) having interhelical interactions which are geometrically similar to those within each monomer[5]. As described in Fig. 2*a*, the symmetrical 4-α-helical packing arrangements which best satisfy the steric requirements of interhelical packing also generate diad axes of symmetry relating surface residues in adjacent helix pairs. As a result, the apoferritin subunits can pack to generate a dimeric structure whose symmetry is similar to that of the monomer.

The truncation of the divergent bundles of the remaining structures precludes the existence of diad symmetry axes on their faces, with a consequent reduction of their aggregate packing symmetry. In cytochrome *c'* for example, the monomers pack about a molecular diad axis which is independent of any symmetry elements of the monomer helix arrangements (Fig. 2).

TMV coat protein differs from the preceding molecules in that it forms aggregates having either 17- or 16.3-fold radial symmetry[6,7] in which each monomer makes different sets of pairwise helical interactions with its neighbours. However, to a good approximation, the interhelical angle between diagonally related parallel helices of the truncated divergent bundle defines the apical angle of a pyramidal prism circumscribing the molecular subunit (Fig. 3). The average observed value for this angle is 22° (180° − 158°, see above), which corresponds closely to the angle subtended by each subunit in both the TMV disk (360°/17 = 21.2°) and helix (360°/16.3 = 22.1°) structures.

To summarize, it has been shown here that the structural similarity among the 4-α-helical proteins can be attributed to what are basically physical requirements of helix packing and chiral connectivity. Consequently, it is not surprising that this structural arrangement should recur throughout the course of evolution, since it evidently possesses considerable scope for functional and superstructural diversity.

We thank Professor P. M. Harrison for providing apoferritin structural data before publication, and D. W. Weatherford for help with the computer graphics work. This work was supported by NIH grants GM21534 and GM25664, the University of Arizona Computer Center, and a Dreyfus Teacher–Scholar Award to F.R.S.

Received 5 February; accepted 13 June 1980.

1. Argos, P., Rossmann, M. G. & Johnson, J. E. *Biochem. biophys. Res. Commun.* **75,** 83–86 (1977).
2. Ward, K. B., Hendrickson, W. A. & Klippenstein, G. L. *Nature* **257,** 818–821 (1975).
3. Hendrickson, W. A., Klippenstein, G. L. & Ward, K. B. *Proc. natn. Acad. Sci. U.S.A.* **72,** 2160–2164 (1975).
4. Stenkamp, R. E., Sieker, L. C., Jensen, L. H. & McQueen, J. E. Jr *Biochemistry* **17,** 2499–2504 (1978).
5. Banyard, S. H., Stammers, D. K. & Harrison, P. M. *Nature* **271,** 282–284 (1978).
6. Bloomer, A. C., Champness, J. N., Bricogne, G., Staden, R. & Klug, A. *Nature* **276,** 362–368 (1978).
7. Stubbs, G., Warren, S. & Holmes, K. *Nature* **267,** 216–221 (1977).
8. Mathews, F. S., Bethge, P. H. & Czerwinski, E. W. *J. biol. Chem.* **254,** 1699–1706 (1979).
9. Weber, P. C. *et al., Nature* **286,** 302–304 (1980).
10. Richmond, T. J. & Richards, F. M. *J. molec. Biol.* **119,** 537–555 (1978).
11. Efimor, A. V. *J. molec. Biol.* **134,** 23–40 (1979).
12. Richards, F. M. *A. Rev. Biophys. Bioengng* **6,** 151–175 (1977).
13. Ptitsyn, O. B. & Rashin, A. A. *Dokl. Acad. Nauk SSSR* **213,** 473–475.
14. Rao, S. T. & Rossmann, M. G. *J. molec. Biol.* **76,** 241–256 (1973).
15. Rossmann, M. G. & Liljas, A. *J. molec. Biol.* **85,** 177–181 (1974).
16. Rossmann, M. G. & Argos, P. *J. molec. Biol.* **105,** 75–95 (1976).
17. Richardson, J. S., Richardson, D. C., Thomas, K. A., Silverton, E. W. & Davies, D. R. *J. molec. Biol.* **102,** 221–235 (1976).
18. Levitt, M. & Chothia, C. *Nature* **261,** 552–557 (1976).
19. Richardson, J. S. *Nature* **268,** 495–500 (1977).
20. Chothia, C. *J. molec. Biol.* **75,** 295–302 (1973).
21. Sternberg, M. J. E. & Thornton, J. M. *J. molec. Biol.* **105,** 367–382 (1976).
22. Richardson, J. S. *Proc. natn. Acad. Sci. U.S.A.* **73,** 2619–2623 (1976).
23. Sternberg, M. J. E. & Thornton, J. M. *J. molec. Biol.* **110,** 269–283 (1977).
24. Weatherford, D. W. & Salemme, F. R. *Proc. natn. Acad. Sci. U.S.A.* **76,** 19–23 (1979).
25. Ramachandran, G. N., Lakshminarayan, A. V. & Kolaskar, A. S. *Biochim. biophys. Acta* **303,** 8–13 (1973).
26. Bernstein, F. C. *et al. J. molec. Biol.* **112,** 535–542 (1977).
27. McLachlan, A. D., Bloomer, A. C. & Butler, P. J. G. *J. molec. Biol.* **136,** 203–224 (1980).
28. Crick, F. H. C. *Acta crystallogr.* **6,** 689–697 (1953).
29. Hendrickson, W. A. & Ward, K. B. *J. biol. Chem.* **252,** 3012–3018 (1977).

Do the chromosomes of the kiwi provide evidence for a monophyletic origin of the ratites?

Leobert E. M. de Boer

Biological Research Department, Royal Rotterdam Zoological and Botanical Gardens, Rotterdam, The Netherlands

The extensive literature on the origin of the ratites focuses mainly on three questions: are the ratites mono- or polyphyletic, did they evolve from flying ancestors, and are they primitive or advanced? Opinion tends to accept a common descent from flying ancestors for the large ratites (for a summary of ideas see ref. 1). They would have evolved on Gondwanaland some time in the Cretaceous and have become dispersed over the southern continents after its fragmentation[2,3]. However, the position of the small New Zealand kiwis, in many respects the most peculiar of all birds, is still a matter for conjecture[1,4]. The chromosome complements of the large ratites have been found to be remarkably uniform[5–8]. The chromosome set of the kiwi, described here, clearly links up with these, which may be recorded as another indication for monophyly of all ratites. It also indicates that we are dealing here with very ancient karyotypes.

0028–0836/80/360084—04$01.00