Volume 250, number 2, 585-590 FEB 07346 July 1989

Isolation of cDNA clones encoding the human Sm B/B′ auto-immune antigen and specifically reacting with human anti-Sm auto-immune sera

N.G. Sharpe, D.G. Williams*, D.N. Howarth, B. Coles+ and D.S. Latchman

Medical Molecular Biology Unit and +Cancer Research Campaign Protein Sequencing Facility, Department of Biochemistry, University College and Middlesex School of Medicine, The Windeyer Building, Cleveland Street, London W1P 6DB and **Division of Clinical Immunology, Mathilda and Terence Kennedy Institute of Rheumatology, London W6 7DW, England*

Received 4 May 1989

A cDNA clone for the human SmB and B′ auto-immune antigens has been isolated by antibody screening of a cDNA expression library. The cDNA clone hybridises with two distinct mRNAs, one of which is expressed in a tissue-specific manner. A fusion protein expressed from the cDNA clone was recognised by a number of sera from systemic lupus erythematosus (SLE) patients containing anti-Sm antibodies but not by sera reactive with other auto-immune antigens. The potential use of this clone in a diagnostic assay for SLE and in elucidating the processes regulating the expression of SmB and B′ is discussed.

Protein, SmB; Protein B′; Autoimmune antigen; Systemic lupus erythematosus; RNA splicing, alternative

1. INTRODUCTION

The sera of some patients with connective tissue diseases contain auto-antibodies directed against a variety of nuclear components [1]. In the case of the auto-immune disease, systemic lupus erythematosus (SLE), certain patients produce antibodies which react with the Sm proteins, two of which, SmB and B′, are similar in size (28 kDa compared to 29 kDa) and closely related to one another [2]. Antibodies to the Sm proteins are restricted to SLE patients and are therefore considered diagnostic for the disease [1].

The availability of anti-sera to these proteins has allowed the demonstration that they and the small nuclear RNAs with which they are associated [3] play an essential role in the splicing of mRNA precursors [4]. In agreement with this, SmB has been detected in a wide range of tissues and cell types [5-7] paralleling the widespread requirement for processing of the primary transcript [8]. In contrast the closely related SmB′ protein is not detectable in most rodent cell types [5-7]. Recently, however, a protein of the size of SmB′ and reacting with antibodies to SmB/B′ has been shown to be expressed in a limited range of rodent cells capable of following an alternative RNA processing pathway [9,10]. In particular such cells are capable of producing the mRNA encoding the calcitonin gene related peptide (CGRP) from the primary transcript of the calcitonin/CGRP gene whilst most cells are capable of splicing this transcript to yield only the mRNA encoding calcitonin [11,12].

In human cells the situation is more complex with all cell types containing both SmB and B′ proteins [5,6]. However, as we have recently shown [9] an acidic isotype of SmB′ is expressed in a tissue specific manner in humans, its expression being confined to the same cell types as those expressing

Correspondence address: N.G. Sharpe, Medical Molecular Biology Unit, Department of Biochemistry, University College and Middlesex School of Medicine, The Windeyer Building, Cleveland Street, London WIP 6DB, England

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/89/$3.50

SmB′ in rodents. Taken together these observations suggest a possible role for this tissue specific Sm protein in the regulation of alternative splicing allowing the production of different mRNAs from the same primary transcript in different cell types.

In order to determine the relationship of this tissue-specific protein to the ubiquitous protein and to investigate whether this Sm protein plays a role in alternative splicing it is necessary to isolate cDNA clones derived from the mRNAs encoding these proteins. The availability of such cDNA clones which expressed the SmB and B′ polypeptides would also allow the detection of antibodies to these proteins in human sera and might therefore be of use in the diagnosis of SLE.

Here we report the use of a monoclonal antibody to isolate such a cDNA clone expressing the SmB/B′ protein as a recombinant fusion protein. This clone has been used to examine the relationship of the tissue specific and ubiquitous forms of these proteins and to monitor the reactivity of human auto-immune sera.

2. MATERIALS AND METHODS

2.1. *Antibody screening and growth of lysogens*

A λgt11 cDNA library prepared from the mRNA of HeLa cells (Clontech) was screened with the KSm5 antibody which reacts with both SmB and B′ [7,9] according to the method of Huynh et al. [13]. The single positive plaque obtained in this procedure was plaque purified by repeated rounds of screening at low density and was used to lysogenise bacteria of *E. coli* strain Y1089 [13]. These bacteria were induced to produce large amounts of recombinant fusion protein by treatment with 10 mM isopropyl-B-D-thiogalacto-pyranoside (IPTG) which is an inducer of the lac operon.

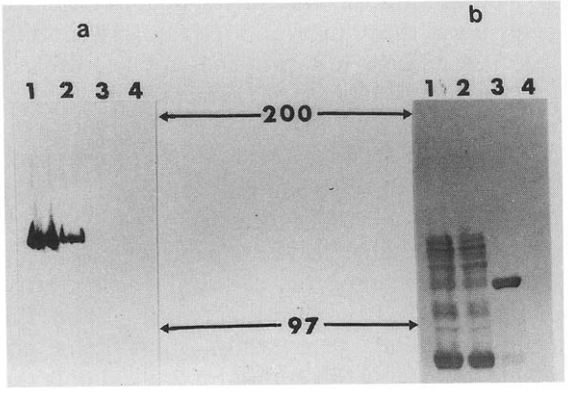

Fig.1. Western blot of protein samples from lysogenic bacterial strains probed with either the KSm5 antibody (panel a) or anti β-galactosidase antibody (panel b). Key to tracks: 1, bacteria lysogenized with K5-1 and induced to produce β-galactosidase by treatment with IPTG; 2, bacteria lysogenized with K5-1 and uninduced; 3, bacteria lysogenized with non-recombinant λgt11 and induced with IPTG; 4, bacteria lysogenized with non-recombinant λgt11 and uninduced. Arrows indicate the positions of marker proteins of the sizes indicated.

2.2. *Western blotting*

Protein isolated from recombinant lysogens was electrophoresed on 5% polyacrylamide-SDS gels [14] and transferred to nitrocellulose by electrophoretic transfer [15]. The nitrocellulose filter was reacted as described [16] with antibody K-Sm5 or human sera as the first layer and peroxidase conjugated anti-mouse or anti-human immunoglobulin as the second layer.

2.3. *Protein purification and sequencing*

SmB protein was purified from rabbit thymus as previously described [17] and the N-terminal sequence determined by automated Edman degradation on an Applied Bio-systems 470A gas phase sequencer with on-line analysis of amino acid-phenylthiohydantoin derivatives. DNA sequence analysis was carried out by the method of Sanger et al. [18] following subcloning of the insert into M13 mp18.

2.4. *RNA analysis*

Cytoplasmic mRNA was isolated by the NP40 lysis method [19] and analysed by Northern blotting as previously described [20]. Blots were hybridised with the 700 base pair insert of K5-1 which had been radio-labelled by the method of Feinberg and Vogelstein [21].

3. RESULTS AND DISCUSSION

To obtain human cDNA clones coding for SmB/B′ a monoclonal anti-Sm antibody, KSm5 [7], was used to screen a HeLa cell cDNA λgt11 expression library. The KSm5 antibody reacts with an immunodominant epitope common to the closely related SmB and B′ proteins [7]. Upon screening 5×10^5 independent λgt11 recombinants, a single positive clone, K5-1, was detected. Isolation of DNA from purified K5-1 yielded an *Eco*RI insert of approx. 700 base pairs.

In the λgt11 expression vector, cDNA clones in the appropriate reading frame will be expressed as a fusion protein with the bacteriophage β-galactosidase protein. To confirm that K5-1 did indeed express such a protein, the recombinant phage was used to lysogenize *E. coli* strain Y1089. After growth of the lysogenic bacteria, they were treated with the lac operon inducer IPTG to induce production of β-galactosidase fusion protein, and the resulting bacterial lysates were analysed by immunoblotting with the KSm5 antibody and an anti

β-galactosidase antibody (fig.1). The bacteria lysogenic for K5-1, as expected, contained a protein reactive with KSm5 which was absent from bacteria lysogenized with non-recombinant vector (fig.1a). This protein was also reactive with anti β-galactosidase antibody (fig.1b) confirming that it was a β-galactosidase–Sm fusion protein.

The fusion protein was approx. 27 kDa larger than native β-galactosidase indicating that K5-1 contains most of the coding region for the SmB/B′ protein. Considerable degradation of the fusion protein was observed in the blot with anti β-galactosidase antibody. This phenomenon which we have previously observed with other recombinant fusion proteins [22] is due to the recognition of such proteins as foreign by the bacteria and results in the rapid removal of the epitope recognized by KSm5 (cf. fig.1a and b) which is located at the extreme carboxyl-terminus of the Sm protein and of the fusion protein (unpublished). In these experiments, no reactivity of the fusion protein was observed with a monoclonal antibody specific for another Sm protein, the 16 kDa SmD protein (not shown).

To prove that the K5-1 cDNA insert represented the protein coding sequence for SmB/B′ rather than being derived from an unrelated protein which fortuitously reacted with KSm5, we obtained the N-terminal amino acid sequence of affinity purified SmB protein (fig.2). Comparison of this sequence with the DNA sequence of the K5-1 insert obtained by the method of Sanger et al. [18] clearly showed a sequence within the K5-1 DNA capable of encoding the SmB protein (fig.2). The open

```
Met Thr Val Gly Lys Ser Ser Lys Met Leu Gln His Ile Asp Tyr Arg Met Arg Cys Ile

.       GAT GGC CGA ATC TTC ATT GGC ACC TTT AAG GCT TTT GAC AAG CAT ATG AAT TTG
Leu Gln Asp Gly Arg Ile Phe Ile Gly Thr Phe Lys Ala Phe Asp Lys His Met Asn Leu

ATC CTC TGT GAT TGT GAT GAG TTC AGA AAG ATC AAG CCA AAG AAT GCG AAG CAA CCA GAG
Ile Leu Cys Asp Cys Asp Glu Phe Arg Lys Ile Lys Pro Lys Asn Ala Lys Gln Pro Glu

CGC GAA GAA AAG CGG GTT TTG GGT CTG GTG TTG CTG CGT GGG GAG AAC TTG GTA TCC ATG
Arg Glu Glu Lys Arg Val Leu Gly Leu Val Leu Leu Arg Gly Glu Asn Leu Val Ser Met

ACT GTG GAG GGG CCA CCC CCC AAA GAT ACT GGC ATT GCT CGG GTA CCA CTT GCT GGA GCT
Thr Val Glu Gly Pro Pro Pro Lys Asp Thr Gly Ile Ala Arg Val Pro Leu Ala Gly Ala

GCT GGA GGC CCT GGG GTT GGT AGG GCA GCT GGT AGA GGA GTA CCA GCT GGT GTG CCA ATT
Ala Gly Gly Pro Gly Val Gly Arg Ala Ala Gly Arg Gly Val Pro Ala Gly Val Pro Ile

CCC CAG GCC CCT GCT GGA TTG GCA GGC CCT GTC CGA GGA GTT GGG GGA CCA TCC CAG CAG
Pro Gln Ala Pro Ala Gly Leu Ala Gly Pro Val Arg Gly Val Gly Gly Pro Ser Gln Gln

GTA ATG ACT CCA CAG GGA AGA GGC ACT GTA GCA GCT GCT GCT GTT GCT GCG ACC GCC AGT
Val Met Thr Pro Gln Gly Arg Gly Thr Val Ala Ala Ala Ala Val Ala Ala Thr Ala Ser

ATT GCT GGA GCC CCA ACA CAG TAC CCA CCA GGA CGG GGC ACT CCG CCC CCA CCC GTC GGC
Ile Ala Gly Ala Pro Thr Gln Tyr Pro Pro Gly Arg Gly Thr Pro Pro Pro Pro Val Gly

AGA GCA ACC CCA CCT CCA GGC ATT ATG GCT CCT CCA CCT GGT ATG AGA CCA CCC ATG GGC
Arg Ala Thr Pro Pro Pro Gly Ile Met Ala Pro Pro Pro Gly Met Arg Pro Pro Met Gly

CCA CCA ATT GGG CTT CCC CCT GCT CGA GGG ACG CCA ATA GGC ATG CCC GCT GCG GGA ATG
Pro Pro Ile Gly Leu Pro Pro Ala Arg Gly Thr Pro Ile Gly Met Pro Pro Pro Gly Met

AGA CCC CCT CCA CCA GGC ATT AGA GGT CCA CCT CCC CCA GGA ATG CGT CCA CCA AGA CCT
Arg Pro Pro Pro Pro Gly Ile Arg Gly Pro Pro Pro Pro Gly Met Arg Pro Pro Arg Pro

TAG CATACTGTTGATCCATCTCAGTCACTTTTTCCCCTGGAATGCGTCTTGTGAAATTG
End
```

Fig.2. DNA sequence of K5-1 and its predicted amino acid sequence. The N-terminal amino-acid sequence of rabbit SmB protein is overlined. Residues 29, 36 and 37 could not be determined unequivocally from the protein sequence and are predicted from the DNA sequence. All other residues in the protein sequence match the predicted amino-acid sequence.

reading frame encoding the appropriate protein sequence continues for a total of 217 amino acids and is followed by 3'-untranslated region indicating that although K5-1 lacks the coding information for the first 23 amino acids of SmB it is capable of encoding most of this protein as predicted from the size of the fusion protein it produces.

Hence K5-1 both reacts with an antibody specific for SmB/B' and contains DNA sequence capable of encoding these proteins. It is therefore derived from an mRNA encoding SmB/B' and can be used to analyze the relationship of these proteins and the processes regulating their expression. In particular we wished to investigate whether the previously observed expression of the SmB' protein in specific rodent cell lines and tissues was reflected at the mRNA level. To do this Northern blot analysis was carried out using a labelled insert from K5-1.

In these experiments (fig.3) a single mRNA of 1.4 kb reacting with the probe was detected in mRNA prepared from mouse liver. Since we have previously shown that this tissue expresses SmB but not SmB' [9], this RNA is likely to encode the SmB protein. In contrast when the K5-1 probe was hybridised to mRNA prepared from an undifferentiated mouse embryonal carcinoma (EC) cell line (PCC3) two mRNA transcripts of approx. 1.6 kb and 1.4 kb were detected the larger of which was present at considerably greater abundance. Since PCC3 cells in common with other EC cells express both the tissue specific SmB' protein and the constitutively synthesized SmB protein [9,10], it is likely that the larger transcript encodes the tissue specific B' protein. The RNA samples gave identical single bands when hybridized with probes specific for other RNA species such as that encoding cytochrome oxidase indicating that the difference in the pattern obtained with the K5-1 probe was not due to RNA degradation (not shown). Hence these results suggest that the tissue and cell specific SmB' protein in rodents is encoded by a distinct tissue-specific mRNA species from that which codes for SmB. The isolation and sequence analysis of cDNA clones representing the two RNA species present in EC cells which we are currently carrying out should allow determination of the relationship between the rodent SmB and B' proteins. Similarly the use of K5-1 to isolate further cDNA clones from the original HeLa cell library will allow an analysis of the differences between the SmB and B' proteins which are constitutively expressed in humans and which are known to be closely related [2].

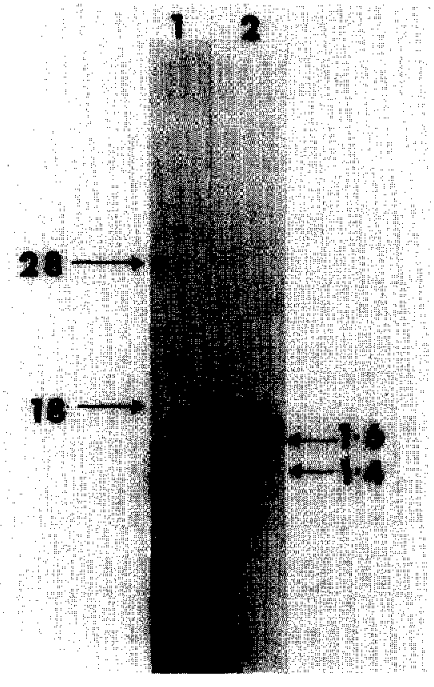

Fig.3. Northern blot of mRNA from mouse liver (track 1) or mouse PCC3 embryonal carcinoma cells (track 2) probed with the insert from K5-1. Arrows indicate the positions of the 28 S and 18 S ribosomal RNAs and of the 1.6 kb and 1.4 kb RNA species detected by the probe.

Ultimately such studies of the SmB/B' family of proteins will pave the way for the elucidation of the role of the tissue specific B' protein in the regulation of alternative splicing. Thus it will be possible to introduce recombinant constructs expressing this protein into cells which normally lack it and to monitor the effect on the ability of such cells to carry out alternative RNA splicing.

In addition to its use in the analysis of the relationships and regulation of the Sm proteins, the fact that K5-1 expresses the SmB/B' protein as a fusion protein with β-galactosidase allowed us to investigate whether such a fusion protein could be used to detect the presence of antibodies to these proteins in human sera. In order to do this, immunoblots of the fusion protein were probed with a range of human sera containing auto-antibodies to different human proteins. The auto-antibody

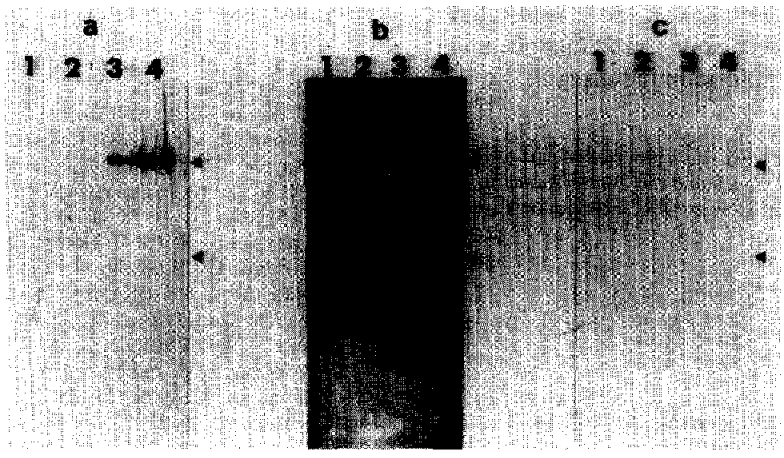

Fig.4. Western blot of protein samples from lysogenic bacteria probed with the KSm5 antibody (panel a), a human serum containing anti-Sm antibodies (panel b) and a human serum lacking anti-Sm antibodies but containing antibodies to other small nuclear RNA-associated proteins (panel c). Key to tracks: 1, bacteria lysogenized with non-recombinant λgt11; 2, as track 1 but induced with IPTG; 3, bacteria lysogenized with K5-1; 4, as track 3 but induced with IPTG. Arrows indicate the positions of native β-galactosidase and of the larger Sm-β-galactosidase fusion protein.

profiles of the sera had been previously determined by immunoblotting with protein extracts prepared from HeLa cells [23]. In these experiments, all five sera tested which exhibited strong reactivity with SmB/B′ in the HeLa cell extract also reacted with the fusion protein. A representative example of the reactivity of such sera is shown in fig.4b. The serum tested exhibits reactivity with the inducible fusion protein in the IPTG-treated sample (fig.4b, track 4) although the strength of the serum is not sufficient for it to detect the low level of fusion protein present in the untreated sample (fig.4b, track 3) which can be detected however, by the higher-affinity monoclonal antibody (fig.4a, track 3). No reactivity of the anti-Sm sera with non-recombinant β-galactosidase was observed in these blots, confirming that they were reacting with the Sm portion of the recombinant fusion protein (see for example, fig.4b, tracks 1 and 2).

To investigate the specificity of this effect we studied the reactivity of the fusion protein with other human sera lacking anti-SmB/B′ antibodies but containing antibodies to other human proteins. In these experiments (fig.4c) the protein failed to react with two anti-nRNP sera containing antibodies to the U1-associated 70 kDa and A proteins which are complexed with the Sm proteins in vivo in a single ribonucleoprotein particle, although these sera readily detected the appropriate proteins in HeLa cell protein samples (not shown). Similarly the fusion protein exhibited no reactivity with sera containing antibodies to other common autoantigens such as SS-A (Ro), or SS-B (La) or the Epstein-Barr virus protein EBNA-1 reactivity to which is commonly present in human sera [1,24].

This indicates that the anti-Sm antibodies in SLE patient sera invariably include species which recognise the Sm epitopes on the fusion protein and that such reactivity is specific to anti-Sm antibodies. Since such antibodies when present in sera, are considered diagnostic for SLE, the immunoreactive SmB/B′ fusion protein could be used for rapid detection of this disease.

Thus the fusion protein might provide a cheap and reliable source of antigen for use in a highly specific diagnostic test for SLE which is currently diagnosed by monitoring reactivity of sera with total rabbit thymus protein [1]. Moreover the absence of any reactivity to non-recombinant β-galactosidase in the sixteen sera tested (see for example fig.4, panels b and c) indicates the possibility of using the fusion protein in an ELISA assay of the type which has been successfully used to detect and quantitate levels of antibodies to the SS-B (La) protein in human sera [25].

In summary therefore the isolation of a cDNA clone derived from an mRNA encoding SmB/B′ should greatly facilitate the study of the relation-

ship of the B and B′ proteins and their role in RNA splicing as well as being of potential diagnostic importance in human auto-immune disease.

Acknowledgements: This work was supported in part by the Arthritis and Rheumatism Council and the Cancer Research Campaign. N.G.S. was supported by an MRC postgraduate studentship.

REFERENCES

[1] Tan, E.M. (1982) Adv. Immunol. 33, 167–240.
[2] Reuter, R., Rothe, S. and Luhrmann, R. (1987) Nucleic Acids Res. 15, 4021–4034.
[3] Bringmann, P. and Luhrmann, R. (1986) EMBO J. 5, 3509–3516.
[4] Maniatis, T. and Reed, R. (1987) Nature 325, 673–678.
[5] Guldner, H.H., Lakomek, H.J. and Bautz, F.A. (1983) J. Immunol. Methods 64, 45–59.
[6] Hintersberg, M., Petterson, I. and Steitz, J.A. (1983) J. Biol. Chem. 258, 2604–2613.
[7] Williams, D.G., Stocks, M.R., Smith, P.R. and Maini, R.N. (1986) Immunology 58, 495–500.
[8] Padgett, R.A., Grabowski, P.J., Konarska, M.M., Seller, S. and Sharp, P.A. (1986) Annu. Rev. Biochem. 55, 1119–1150.
[9] Sharpe, N.G., Williams, D.G., Norton, P. and Latchman, D.S. (1989) FEBS Lett. 243, 132–136.
[10] McAllister, G., Amara, S.G. and Lerner, M.R. (1988) Proc. Natl. Acad. Sci. USA 85, 5296–5300.
[11] Leff, S.E., Evans, R.M. and Rosenfeld, M.G. (1987) Cell 48, 517–524.
[12] Crenshaw, E.B., Russo, A.F., Swanson, L.W. and Rosenfeld, M.G. (1987) Cell 48, 517–524.
[13] Huynh, T.V., Young, R.A. and Davis, R.W. (1985) in: DNA Cloning Techniques: A Practical Approach (Glover, D.M. ed.) pp. 49–78, IRL Oxford.
[14] Laemmli, U.K. (1970) Nature 227, 680–685.
[15] Towbin, H., Staehelin, T. and Gordon, J. (1979) Proc. Natl. Acad. Sci. USA 76, 5350–5354.
[16] La Thangue, N.B., Shriver, K., Dawson, C. and Chan, W.L. (1984) EMBO J. 3, 267–277.
[17] Williams, D.G., Charles, P.J. and Maini, R.N. (1988) J. Immunol. Methods 113, 25–35.
[18] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[19] Favoloro, J., Treisman, R. and Kamen, R. (1980) Methods Enzymol. 65, 718–749.
[20] Murphy, D., Brickell, P.M., Latchman, D.S., Willison, K. and Rigby, P.W.J. (1983) Cell 35, 865–871.
[21] Feinberg, A.P. and Vogelstein, B. (1983) Anal. Biochem. 132, 6–13.
[22] Patel, R., Chan, W.L., Kemp, L.M., La Thangue, N.B. and Latchman, D.S. (1986) Nucleic Acids Res. 14, 5629–5640.
[23] Williams, D.J., Stocks, M.R., Charles, P.J. and Maini, R.N. (1986) J. Immunol. Methods 91, 65–73.
[24] Morrow, J. and Isenberg, D. (1987) Auto-Immune Rheumatic Disease, Blackwell Scientific, Oxford.
[25] St. Clair, E.W., Pisetsky, D.S., Reich, C.F., Chambers, J.C. and Keene, J.D. (1988) Arth. Rheum. 31, 506–514.