ARTICLE

Streptomyces aureorectus DSM 41692 and *Streptomyces virens* DSM 41465 are producers of the antibiotic nucleocidin and 4'-fluoroadenosine is identified as a co-product

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Yawen Chen,[a] Qingzhi Zhang,[a] Xuan Feng[a], Marta Wojnowska[a] and David O'Hagan*[a]

Abstract: Genome homology and the presence of a putative biosynthetic gene cluster identified *Streptomyces aureorectus* DSM 41692 and *Streptomyces virens* DSM 41465 as candidate producers of the antibiotic nucleocidin **1**. Indeed when these bacterial strains were cultured in a medium supplemented with fluoride (4mM) they each produced nucleocidin **1** and the previously identified 4'-fluoro-3'-O-*β*–glucosylated adenosine **2** and its sulfamylated derivative **3**. In both of these cases 4'-fluoroadenosine **9** is also identified as a natural product although it has never been observed during fermentations of *Streptomyces calvus,* the original source of nucleocidin **1**. The identity of 4'-fluoroadenosine **9** was confirmed by a total synthesis as well as by its *in vitro* enzymatic conversion to metabolite **2** using the glucosyl transferase enzyme, NucGT.

Introduction

Nucleocidin **1**[1] is among the rare fluorine containing metabolites.[2] The structure of nucleocidin is sufficiently unique that the producing strains must contain a fluorination enzyme that differs from the fluorinase involved in fluoroacetate biosynthesis, the latter being the only C-F bond forming enzyme that has been characterised so far.[2b,3] The search for a novel fluorination enzyme has piqued a recent interest in nucleocidin biosynthesis.[4-6] Nucleocidin **1** also possesses an unusual sulfamyl ester moiety, a structural motif restricted so far to nucleocidin and its co-metabolites **2** and **3** as well as the closely related ascamycin antibiotics **6** and **7**.[7] Unlike nucleocidin **1** the ascamycins **6** and **7** have a 2-chloroadenosine ring, but they do not possess a ribosyl fluorine. Nucleocidin **1** was originally identified as a metabolite of the soil bacterium *Streptomyces calvus,*[1] however available strains appeared to have lost the ability to produce the metabolite in culture and this hampered biosynthetic studies for many years. More recently full genome sequence analysis of *S. calvus* strains in the public domain revealed bald-gene (*bld*) mutations that, when corrected, re-established nucleocidin production and this has led to a revival in the interest in nucleocidin biosynthesis.[4b] An industrial strain without the mutation, *S. calvus* T-3018, also produces the antibiotic in culture.[5] Analysis of the co-metabolites from *S. calvus* T-3018 allowed the identification of the two 4'-fluoro-3'-O-*β*–glucosylated adenosines **2** and **3** which are found only in the supernatant, presumably exported from the cell after 3'-OH *β*-glucosylation.[5b] Interestingly, the corresponding metabolites **4**, **5** and **8**, without fluorine at C-4' of the ribose, are also co-metabolites in extracts of *S. calvus* T-3018.[5a] In addition we have identified a glucosyl transferase (NucGT) which, when over-expressed, has the ability to catalyse *in vitro β*-glucosylation of the 3'-OH group of the ribose ring of nucleocidin **1** with UDP-glucose, generating metabolite **3**, as well as the related biotransformations to the nonfluorinated metabolites **4** and **5**.[5a,5b] The *nucGT* gene is close to a cluster of genes which appear by annotation to be involved in sulfamylation. This gene cluster is homologous to that implicated in ascamycin **6** and dealanyl-ascamycin **7** production by *Streptomyces sp.* JCM9888,[8] further supporting a role in sulfamyl ester antibiotic biosynthesis. These gene clusters and their common features are depicted in Fig 2 and are highlighted in detail in the SI. Most recently genome mining led to the identification of *Streptomyces asterosporus* DSM 41452 as a potential nucleocidin producer.[4a] Genetic manipulation was required to correct a rare codon usage mutation, after which *S. asterosporus* DSM 41452 was also confirmed as a nucleocidin **1** producer. Therefore, to date two bacterial strains have been identified as producers of nucleocidin **1** and its co-metabolites **2** and **3**. In this work we identify a further two *Streptomyces* strains based on genome sequence homologies which we show are producers of nucleocidin **1** and co-metabolites **2** and **3**. We also identify 4'-fluoroadenosine **9** from these two strains, a metabolite that does not appear to accumulate in *S. calvus* or *S. asterosporus* DSM 41452, and as such constitutes a new fluorometabolite.

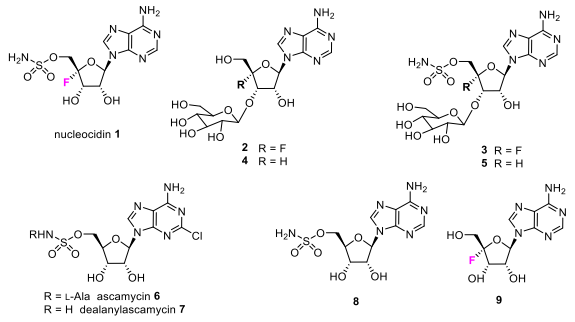


Figure 1. Sulfamyl adenosine antibiotics and related metabolites **1 – 8**.

Results and discussion

Using individual genes within the putative nucleocidin biosynthetic gene cluster of *S. calvus* T-3018 as search tools against those in the public domain, led to identification of analogous gene clusters in *Streptomyces aureorectus* DSM 41692 and *Streptomyces virens* DSM 41465.[10,11] Like *Streptomyces asterosporus* DSM 41452, which is the only other known nucleocidin producer,[4a] *S. aureorectus* DSM 41692 and *S. virens* DSM 41465 appear to have almost identical gene clusters to the *S. calvus* one,[4a] while *S. sulfonofaciens* DMS 41679[12]

[a] *School of Chemistry, University of St Andrews, North Haugh, St Andrews, Fife, KY16 9ST, UK. E-mail: do1@st-andrews.ac.uk*

† Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

possesses a similar, though not identical, gene cluster, with >70% sequence identity (Fig. 2). To establish the relatedness of these strains their 16S rRNA sequences were compared, with *S. calvus* T-3018 16S rRNA showing > 99% sequence identity to *S. asterosporus* DSM 41452, *S. aureorectus* DSM 41692 and *S. virens* DSM 41465, and 96.6% to *S. sulfonofaciens* DSM 41679 (Fig. 3).

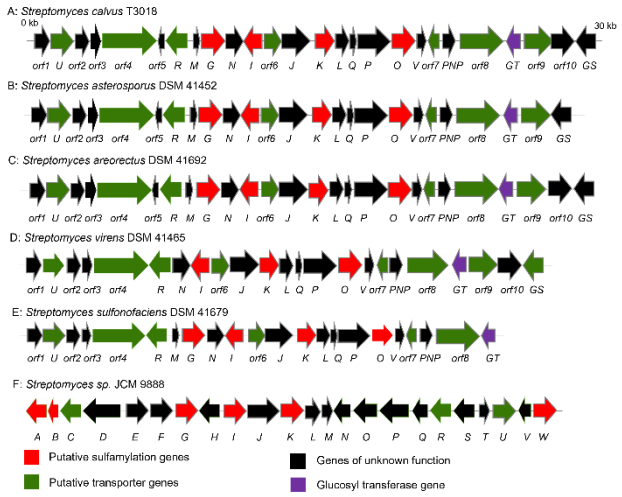


Figure 2. Schematic representation of the putative gene clusters associated with nucleocidin **1** and ascamycin **6** (*Streptomyces sp*. JCM 988) biosynthesis in various *Streptomyces*. More detailed annotations are reported in the SI.

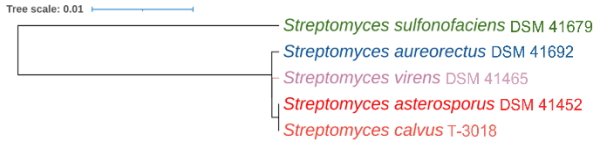


Figure 3. Phylogenetic tree based on the 16S rRNA gene sequences. The tree was generated by iTOL.[14]

Given these close genetic similarities, strains of *S. aureorectus* DSM 41692, *S. virens* DSM 41465 and *S. sulfonofaciens* DSM 41679 were obtained from the DSMZ culture collection and each was cultured in the laboratory and assessed for fluorometabolite production.

In all three cases the cultures were grown in the recommended production medium supplemented with 4 mM fluoride ion. Nucleocidin **1** and associated fluorometabolites **2** and **3** were obvious by $^{19}F\{^{1}H\}$-NMR in extracts of *S. aureorectus* DSM 41692 (Fig 4) and *S. virens* DSM 41465 (Fig 5), however there was no detectable production of any fluorocompounds from *S. sulfonofaciens* DSM 41679 (not shown). In the producing organisms, cells were harvested between days 4-8 and the fluorometabolites were extracted into

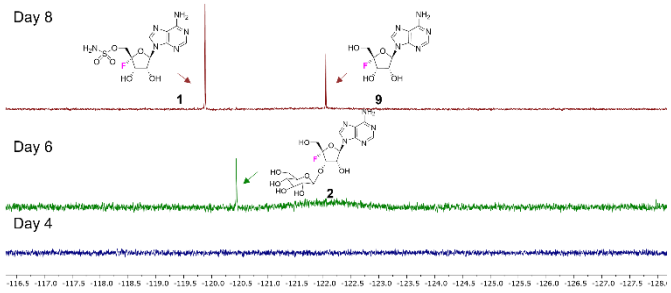


Figure 4. $^{19}F\{^{1}H\}$-NMR time course of fluorometabolites production in *S. aureorectus* DMS 41692 over 8 days.

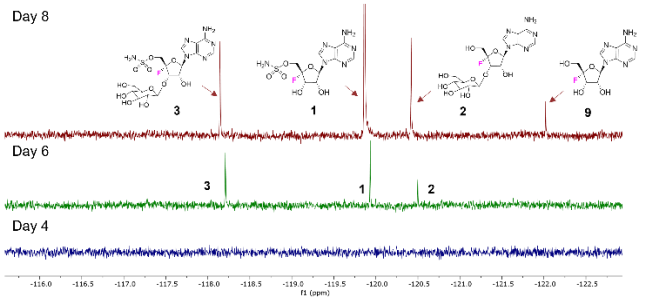
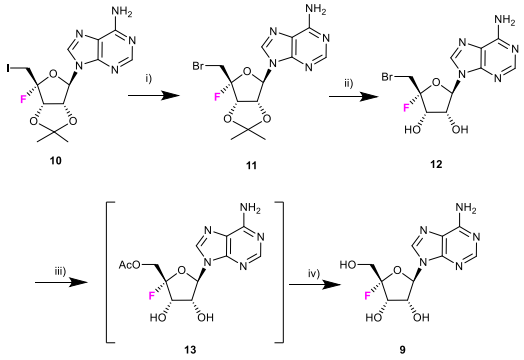


Figure 5. $^{19}F\{^{1}H\}$-NMR time course of fluorometabolites production in *S. virens* over 8 days.

butanol as previously described.[5b] Time course analysis identified metabolites **2** and **3** as the first products to form with nucleocidin **1** accumulating and becoming the dominant metabolite around days 6-8. In extracts of both of these strains a fourth organofluorine metabolite was evident which appeared quite late in the production profile, after the disappearance of co-metabolites **2** and **3**. This novel compound is visible in the $^{19}F\{^{1}H\}$-NMR spectra of the day 8 extracts in both Fig 4 and Fig 5. We have never observed this metabolite in our extensive culturing of *S. calvus* T-3018 and it has not been reported elsewhere. Attempts at isolating it by purification failed, implying that it is unstable. We noticed that the ^{19}F-NMR chemical shift of -122 ppm is identical to that of 4'-fluoroadenosine **9**, a compound that has previously been prepared synthetically.[15] The two reported syntheses [15,16] of 4'-fluoroadenosine **9** indicated it to be unstable, with a half-life that varies from minutes to hours depending on the pH, and this lack of stability is consistent with our experience in isolating the unknown metabolite. Therefore, it became an objective to establish if the new metabolite was indeed **9**, and two approaches were taken. The first involved preparing 4'-fluoroadenosine **9** by synthesis to use as a reference compound and then adding it to the extract to supplement the $^{19}F\{^{1}H\}$-NMR signal. The other approach involved incubating an extract from day 8 with the glucosyl transferase enzyme NucGT from *S. calvus,* in an assay with UDP-glucose.[5b] This should convert 4'-fluoroadenosine **9** to known metabolite **2** and thus establish its identity by proxy.



Scheme 1. Synthetic route to 4'-fluoroadenosine **9**. Reagents and conditions: i) NaBr, DMF, 97°C, 24 h, 51%); ii) 60% aqTFA, 0°C to rt, 64%; iii) KOAc, DMF, 97°C, 48 h; iv) Na_2CO_3 , MeOD, rt, 30 min; 20% yield over the last two steps.

Firstly 4'-fluoroadenosine **9** was prepared as illustrated in Scheme 1. The route started from 5'-iodo-4'-fluoroadenosine acetonide **10**, an intermediate reported previously in two independent total syntheses of nucleocidin **1**.[16] Attempts to release the acetonide moiety from **10** with TFA or 4N HCl in dioxane led to its complete decomposition. However, after exchanging the iodine in **10** for bromine in a Finkelstein protocol, the acetonide of 5'-bromo-4'-fluoroadenosine acetonide **11** was safely removed with 60% aq TFA. This gave 5'-

bromo-4'-fluoroadenosine **12** without significant decomposition. Nucleophilic substitution of the 5'-bromine of **12** with potassium acetate in heated DMF gave 5'-acetyl 4'-fluoroadenosine **13** along with unreacted **12**, and a minor amount of target **9**. However, after replacing DMF with methanol as the solvent, treatment of this mixture with sodium carbonate resulted in methanolysis of the acetyl group of **13**, to achieve a full conversion to 4'-fluoroadenosine **9**. A similar approach was reported for the last step of a previous preparation of **9** from 2',3',5'-tribenzoyl adenosine.[15a] The product was then purified by semi-preparative HPLC to afford **9**. As previously reported,[14] this is an unstable compound and was kept at -80 °C until required for analytical studies. Addition of synthetic **9** to an extract of *S. aureorectus* DSM 41692 which contained nucleocidin **1** and the unknown compound, in an add-mix experiment, showed that the $^{19}F\{^{1}H\}$-NMR signal at -121.6ppm increased in intensity, consistent with the unknown metabolite being 4'-fluoroadenosine **9**. This experiment is illustrated in Fig 6.

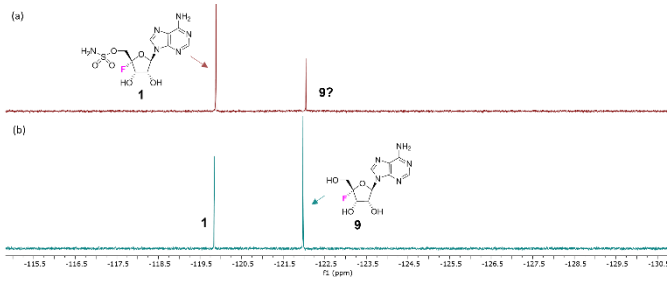


Figure 6. Profiles of an add-mix experiment of natural and synthetic 4'-fluoroadenosine **9**. The upper $^{19}F\{^{1}H\}$-NMR spectrum (a) is that from a butanol extract of *S. aureorectus* DSM 41692 on day 8. The lower $^{19}F\{^{1}H\}$-NMR spectrum (b) is recorded after the addition of synthetic 4'-fluoroadenosine **9** to the above extract.

The identity of 4'-fluoroadenosine **9** as a metabolite was further confirmed by its enzymatic conversion to co-metabolite **2**. An extract of *S. aureorectus* DSM 41465 which contained 4'-fluoroadenosine **9** (and nucleocidin **1**), was incubated with over-expressed NucGT and UDP-glucose,[5b] and there was an immediate conversion to **2** (and **3**) as determined by $^{19}F\{^{1}H\}$-NMR (Fig 7b) although the conversion of **9** to **2** (relative to that from **1** to **3**) is reduced presumably because **9** is an unstable compound.

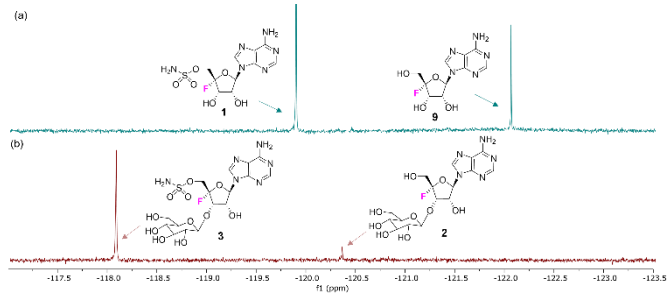


Figure 7. (a) $^{19}F\{^{1}H\}$-NMR spectrum of extract of *S. aureorectus* DSM 41692 containing **1** and **9**; (b) $^{19}F\{^{1}H\}$-NMR spectrum of the extract from (a) of *S. aureorectus* DSM 41692 after incubation with NucGT and UDP-glucose.

The glucosylation of synthetic 4'-fluoroadenosine **9** with NucGT was also investigated. NucGT assays were performed in Tris-HCl buffer (50 mM, pH 8.0) with UDP-glucose (4.4 mM) and $MgCl_2$ (25 mM). The mixture was incubated for 5 min at room temperature prior to $^{19}F\{^{1}H\}$-NMR analysis; the resulting spectra indicated the conversion of synthetic **9** to metabolite **2** as illustrated in Fig 8. These NucGT assays further confirm that 4'-fluoroadenosine **9** is a natural product of both *S. aureorectus* DSM 41692 and, by implication, *S. virens* DSM 41465.

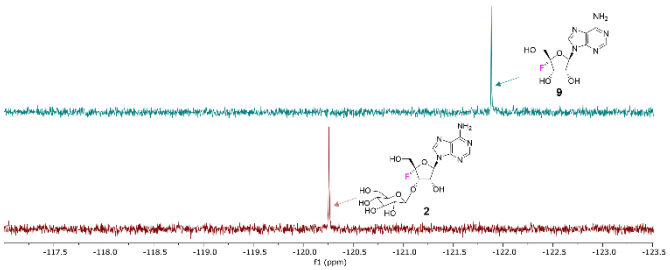


Figure 8. $^{19}F\{^{1}H\}$-NMR spectrum (upper) of synthetic fluoroadenosine **9** and the $^{19}F\{^{1}H\}$-NMR spectrum (lower) of synthetic fluoroadenosine **9** after incubation with NucGT and UDP-glucose, indicating conversion to fluorometabolite **2**.

Conclusions

In conclusion, we have searched the public domain for homologues of the putative nucleocidin biosynthetic gene cluster of *S. calvus* T-3018 and identified three candidate nucleocidin **1** producers. Two of these were demonstrated to produce nucleocidin **1** and its co-metabolites **2** and **3**, in a similar manner to *S. calvus* T-3018 and *S. asterosporus* DSM 41452, the only other known producer of the antibiotic. This brings the number of nucleocidin producing strains to four. The two strains reported here, *S. aureorectus* DSM 41692 and *S. virens* DSM 41465 also produce 4'-fluoroadenosine **9**, extending the range of the rare fluorometabolites. Its appearance relatively late in the fluorometabolite production profile suggests that 4'-fluoroadenosine **9** is a metabolite of nucleocidin **1** or co-metabolites **2** or **3**, rather than an early biosynthetic intermediate of the fluorometabolite pathway, although the details of this pathway remain yet to be elucidated.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgements

We thank the Leverhulme Trust, EPSRC and the Sinfonia consortium of the EU for grants as well as the Chinese Scholarship Council for a Studentship (YC), all of which helped to support this work.

Notes and references

1. (a) S. O. Thomas, V. L. Singleton, J. A. Lowery, R. W. Sharpe, L. M. Pruess, J. N.Porter, J. H. Mowat, N. Bohonos, *Antibiot. Annu.,* 1956, 1956-1957; (b) I. D. Jenkins, J. P. H. Verheyden, J. G. Moffatt, *J. Am. Chem. Soc.,* 1976, **98**, 3346.
2. (a) M. F. Carvalho, R. S. Oliveira, *Crit. Rev. Biotech.,* 2017, **37**, 880-897; (b) D. O'Hagan, H. Deng, *Chem. Rev.,* 2015, **115**, 634-649.
3. (a) H. M. Senn, D. O'Hagan, W. Thiel, *J. Am. Chem. Soc.,* 2005, **127**, 13643-13655; (b) K. K. Chan, D. O'Hagan, *Meth. Enzymol*., 2012, **516**, 219-235.
4. (a) S. Zhang, D. Klementz, J. Zhu, R. Makirtynskyy, A. R. O. Pasternak, S. Günther, D. L. Zechel, A. Bechthold, *J. Biotechnol.*, 2019, **292**, 23 – 31; (b) X. M. Zhu, S. Hackl, M. N. Thaker, L. Kalan, C. Weber, D. S. Urgast, E. M. Krupp, A. Brewer, S. Vanner, A. Szawiola, G. Yim, J. Feldmann, A. Bechthold, G. D. Wright, D. L. Zechel, *ChemBioChem.,* 2015, **16**, 2498 – 2506; (c) L. Kalan, A. Gessner, M. N. Thaker, N. Waglechner, X-M, Zhu, A. Szawiola, A. Bechthold, G. D. Wright, D. L. Zechel, *Chem. Biol,* 2013, **20**, 1214 - 1224
5. (a) X. Feng, D. Bello, D. O'Hagan, *RSC Advances*, 2021, **11**, 5291–5294; (b) X. Feng, D. Bello, P.T. Lowe, J. Clark, D. O'Hagan, *Chem Sci.,* 2019, **10**, 9501; (c) X. Feng, N. Al Maharik, A. Bartholomé, J. E. Janso, U. Reilly, D. O'Hagan, *Org. Biol. Chem*., 2017, **15**, 8006–8008; (d) A. Bartholomé, J.

E. Janso, U. Reilly, D. O'Hagan, *Org. Biol. Chem.,* 2017, **15**, 61 - 64.
6. U. Ngivprom, S. Kluaiphanngam, W. Ji, S. Siriwibool, A. Kamkaew, J. R. K. Cairns, Q. Zhang, R-Y Lai, *RSC Adv.,* 2021, **11**, 3510-3515.
7. (a) P. Mujumdar, S. Bua, C. T. Supuran, T. S. Peat, S-A. Poulsen, *Bioorg. Med. Chem. Letters,* 2018, **28**, 3009-3013; (b) K. Isono, M. Uramoto, H. Kusakabe, N. Miyata, T. Koyama, M. Ubukata, S. K. Sethi, J. A. McCloskey *J. Antibiot*., 1984, **37**, 670 - 672. (c) E. Takahashi, T. Beppu, *J. Antibiot*., 1982, **35**, 939 - 947.
8. (a) C. Zhao, J. Qi, W. Tao, L. He, W. Xu, J. Chen and Z. Deng, *PLoS One*, 2014, **9**, e114722; (b) K. Isono, M. Uramoto, H. Kusakabe, N. Miyata, T. Koyama, M. Ubukata, S. K. Sethi and J. A. McCloskey, *J. Antibiot.,* 1984, **37**, 670–672.
9. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, *J. Mol. Biol.,* 1990, **215**, 403 - 410.
10. K. Li., S. Hu., Y. Wang., Y. Guo., M. Zhou., X. Tang., J. Gao., *Int. J. Syst. Evol. Microbiol.,* 2021, **71**, 004955.
11. G. F. Gauze, M. A. Sveshnikova, T. S. Maksimova, O. L. Olkhovatova, V. S. Bazhanova, *Antibiotiki,* 1977, **22**, 963-967.
12. S. Miyadoh, T. Shomura, T. Ito, T. Niida, *J. Int. J. Syst. Bacteriol.,* 1983, **33**, 321 - 324.
13. E. Stackebrandt, D. Witt., C. Kemmerling, R. Kroppenstedt, W. Liesack, *Appl. Environ. Microbiol.,* 1991, **57**, 1468 - 1477.
14. I. Letunic, P. Bork, *Nucleic Acids Res*., 2021, **49**, 293-296.
15. a) S. Lee, C. Uttamapinant, G. L. Verdine, *Org. Lett.,* 2007, **9**, 24, 5007-5009; (b) D. Guillerm, M. Muzard, B. Allart, G. Guillerm, G. *Bioorg. Med. Chem. Lett.,*1995, **5**, 1455- 1460.
16. (a) A. R. Maguire, W. Meng, S. M. Roberts, A. J. Willets, *J. Chem. Soc. Perkin Trans I*, 1993, **15**, 1795-1808; (b) I. D. Jenkins, J. P. H. Verheyden, J. G. Moffatt, *J. Am. Chem. Soc.,* 1976, **98**, 3346 -3357.