Organic & Biomolecular Chemistry

The secondary structure of a heptapeptide containing trifluoromethyl-λ^6-tetrafluorosulfanyl substituted amino acids on

Journal:	*Organic & Biomolecular Chemistry*
Manuscript ID	OB-COM-08-2019-001797.R1
Article Type:	Communication
Date Submitted by the Author:	19-Aug-2019
Complete List of Authors:	Ikeda, Akari; University at Albany, Chemistry Capellan, Aimée ; University at Albany Welch, John; University at Albany, Chemistry

Organic & Biomolecular Chemistry

COMMUNICATION

The secondary structure of a heptapeptide containing trifluoromethyl-λ⁶-tetrafluorosulfanyl substituted amino acids

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

DOI: 10.1039/x0xx00000x

www.rsc.org/

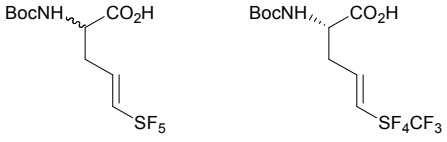
Akari Ikeda,[a] Aimée Capellan[a] and John T. Welch*[a]

Abstract. Site specific introduction of the polar hydrophobic trifluoromethyl-λ⁶-tetrafluorosulfanyl (CF_3SF_4) group can effectively control the secondary structure of a heptapeptide, the minimum repeat unit of an α-helix. The structural inflluence of CF_3SF_4-containing amino acid on the heptapeptide was established using NMR methods.

The reactivity and conformation of amino acids is often influenced by fluorination. The most common fluorinated amino acids, leucine and valine, typically contain three to six omega fluorines commonly in trifluoromethyl (CF_3) groups. In addition to well-known lipophobic[1] effects or specific fluorine-fluorine interactions,[1a, 2] CF_3 groups may have other structural influences. In particular, fluorination[3] may affect polypeptide secondary structure, such as helix-formation.[1b, 2a, 2b, 3g, 4] Incorporation of fluorinated amino acids may result in a reduced propensity for α-helix formation relative to native amino acid-containing sequences.[5] The accessibility of various secondary structures[4d, 6] may be diminished by electrostatic interactions of fluorinated side chains with the amide backbone. However, incorporation of fluorinated analogues of leucine, isoleucine or valine at the core of coiled-coil heptad repeats can result in enhancement of the thermal stability of coiled-coil peptide assemblies.[2a, 4d, 7]

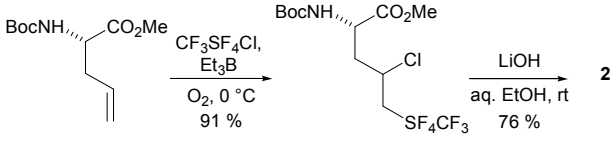
The simple and convenient preparation of *trans*-trifluoromethyl-λ⁶-tetrafluorosulfanyl chloride facilitated the introduction of the CF_3SF_4 group[8] and hence the preparation and resolution of CF_3SF_4-substituted amino acids. The CF_3SF_4 group is one of the most hydrophobic groups known with an experimentally determined lipophilicity partition coefficient of 2.13[9] (π_p), significantly greater that that of the pentafluorosulfanyl group (SF_5) (1.23)[9] or the trifluoromethylsulfanyl group (SCF_3) (1.44). The large lipophilicity increments (π_p) and high Hammett substituent values (σ_p) associated with CF_3SF_4-substitution are associated with decreased desolvation energy and enhanced dipolar interactions[10] the same properties correlated with polar hydrophobicity.[11] The Connolly volume, 138.93 Å^3, and surface area,[9] 156.76 Å^2, of the CF_3SF_4 group are also greater than the corresponding values of the pentafluorosulfanyl (SF_5) group, 102.96 Å^3 and 122.71 Å^2 respectively. When the CF_3SF_4 group is introduced to the first and fifth residues of a heptapeptide, the substituent effects were predicted to influence the heptapeptide secondary structure. NMR methods demonstrated that even the sterically demanding CF_3SF_4-substituted sidechains of both the protected and deprotected peptides strongly associate in an intramolecular manner.

Earlier the influence of the racemic pentafluorosulfanylated amino acid (*RS*)-**1** on the conformation[12] of a heptapeptide substituted with **1** at the first and fifth positions was examined. NMR solution structure determination methods[13] enabled the tentative assignment of four coiled structures that were tentatively assigned to the four diastereomeric heptapeptides that were prepared.[12] In marked contrast to readily available aryl pentafluorosulfanyl compounds,[14] aliphatic SF_5-containing building blocks or pentafluorosulfanyl halides are inaccessible. This inaccessibility made the preparation of optically pure **1,** or other aliphatic pentafluorosulfanylated amino acids, economically challenging.



1 **2**

In contrast the ready availability of CF_3SF_4Cl made testing the efficacy of CF_3SF_4-substitution on controlling secondary peptide



3 **4**

Scheme 1. Preparation of (*S,E*)-*N*-(*tert*-butoxycarbonyl)-2-amino-6-trifluoromethyl-λ⁶-tetrafluorosulfanylpent-4-enoate(*Boc*-TtsNVa) **2.**

structure much more accessible. (*S,E*)-*N*-(*tert*-butoxycarbonyl)-2-amino-6-trifluoromethyl-λ⁶-tetrafluoro-sulfanylpent-4-enoic acid

[a] *Department of Chemistry, University at Albany, SUNY, 1400 Washington Ave., Albany NY 12222.*

Electronic Supplementary Information (ESI) available: Experimental section and spectroscopy. See DOI: 10.1039/x0xx00000x

(Boc-TtsNVa) **2** was prepared by the addition of CF_3SF_4Cl[8] to methyl (*S*)-*N*-(*tert*-butoxycarbonyl)-2-aminopent-4-enoate **3**. Dehydro chlorination and saponification formed **2**(Scheme 1) Incorporation of TtsNVa at the first and fifth position of a heptapeptide such as **5** (TtsNVa-Glu-Ser-Lys-TtsNVa-Lys-Glu or TtsNV-E-S-K-TtsNV-K-E) (Fig. 1) was anticipated to promote a helical conformation**.**

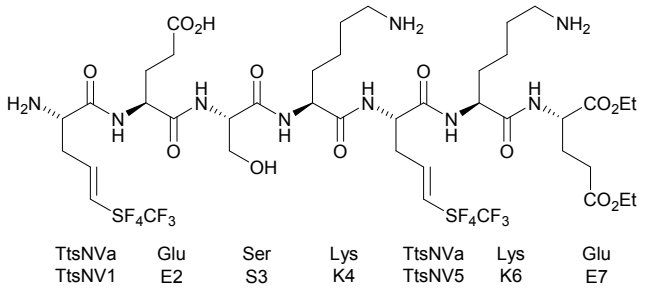


Figure 1. The CF_3SF_4--containing heptapeptide **5** (TtsNVa-Glu-Ser-Lys-TtsNVa-Lys-Glu or TtsNV-E-S-K-TtsNV-K-E).

As illustrated in a helix wheel depiction, the CF_3SF_4-containing amino acids would be adjacent on peptide folding (Fig. 2) .

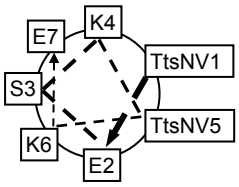
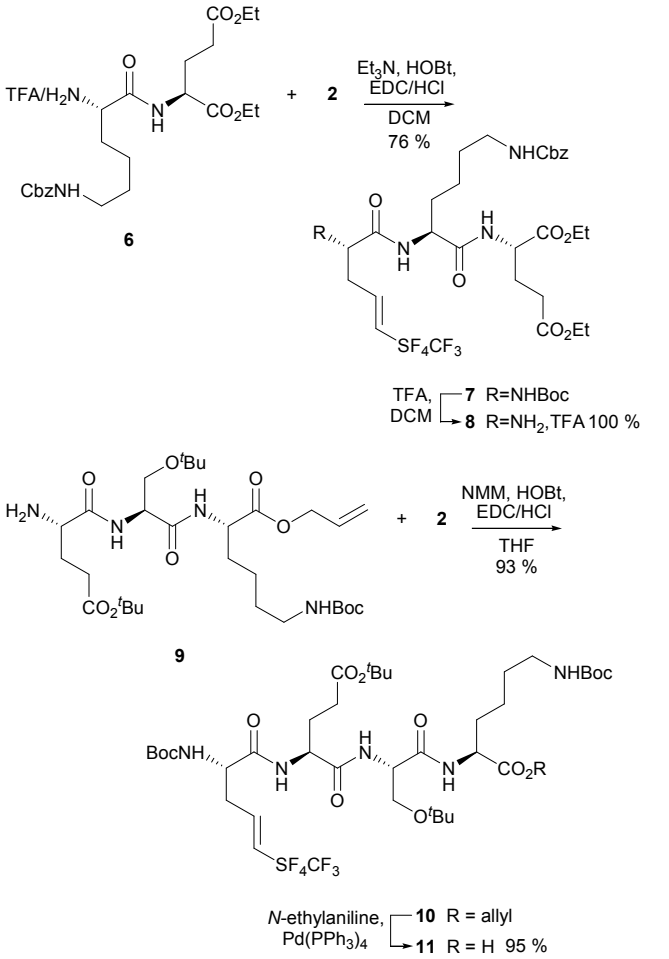


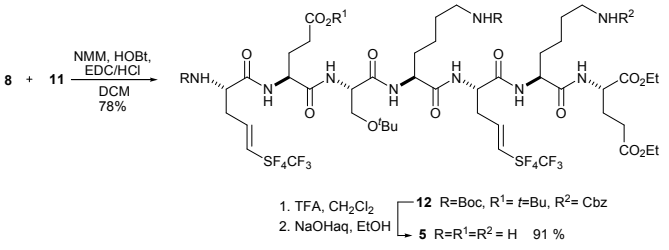
Figure 2. A helix wheel representation of the heptapeptide target showing the CF_3SF_4 groups in close proximity.

C- and *N*-terminal, tri- and tetrapeptide, fragments were prepared as shown (Scheme 2). Prior to purification, **8** was formed in 84% ee as determined by the method of Mosher,[15] i.e, preparation of a α-methoxy-α-trifluoromethylphenyl-acetamide. The minor racemization that occurred on peptide coupling was easily overcome by chromatographic purification of **7**. Addition of **2** to **9** proceeded efficiently and **10** was likewise purified by chromatography. The coupling of **8** with **11** failed under the conditions used for condensation of **2** with **6** or **2** with **9** with only unreacted **8** and **11** recovered. Apparently solvation of the CF_3SF_4-containing peptides strongly influenced the success of the coupling reaction. The optimum reaction conditions required dichloromethane as solvent to disrupt aggregation of the reactant peptides (Scheme 3) .

The influence of the intermolecular CF_3SF_4-group interactions on the coupling reaction suggests that CF_3SF_4-substituents in both the protected and deprotected peptides will affect the secondary structure of both peptides **12** and **5**. The three dimensional structure of the protected peptide **12** was determined by ^{1}H-NMR spectroscopy. One dimensional ^{1}H and ^{19}F-NMR spectra and two dimensional HSQC, ^{1}H,^{1}H ROESY, and ^{1}H,^{1}H TOCSY spectra[16] of the peptide were acquired in $CDCl_3$. The TOCSY data were used to assign the ^{1}H resonances using heteronuclear coupling of protons to the fluorines of CF_3SF_4-group as a point of origin. The ROESY spectrum was used to determine the NOE (nuclear Overhauser effect) cross-peak assignments and to confirm unambiguously the distance relationships between the individual protons (See Supplementary information S2-4-2c, pS24. See S2-4-1d, p S21 for **12**).



Scheme 2. Preparation of *N*- and *C*-terminal fragments, **8** and **11**, by the addition of (TtsNVa) **2** to the corresponding di- and tripeptide fragments .



Scheme 3. Condensation of *N- or C*-terminal fragments to form the targeted heptapeptide **5**.

CARA (Computer aided resonance assignment, http://cara.nmr.ch/doku.php) was used to generate CYANA input from the ROESY data.[13, 17] As **12** is composed of protected natural amino acids and the synthetic amino acid **2,** for the structure fitting, new residue types were constructed in CARA, for inclusion of the additional spins of the protecting

groups and of **2**. This additional information enabled more accurate structure determination, such as sidechain orientation, than was previously possible when racemic pentafluorosulfanylated amino acids were employed.[12] Structural calculations were carried out with the integrated autoassignment module CANDID[13, 17-18] using distance restraints derived from the NOESY spectra and seven pairs of backbone torsion angle constraints derived from TALOS.[19]

The chemical shift identification tolerance was set to 0.020 ppm in both proton dimensions. Initial NOE cross-peaks were manually assigned in the NMR view of CARA. A total of 300 structures were generated per CANDID cycle with 10,000 torsion angle dynamics (TAD) steps in the CYANA annealing protocol. A total of **265** manually and CANDID-assigned NOE cross-peaks were used in structural determination. The 20 lowest-energy structures in the final CANDID round were retained as an ensemble representation of the derived structure.

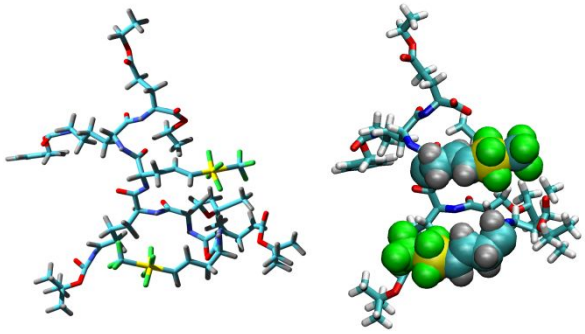


Figure 3. a. The NMR derived structure of the protected peptide **12** in $CDCl_3$. b. The lowest energy conformer of the best 20 structures is depicted with the atoms of the TtsNVa side chain represented as van der Waals spheres.

The most stable of the final 20 structures determined from the $CDCl_3$ solution of **12** illustrates the close proximity of the CF_3SF_4-containing side chains of the TtsNVa residues (Fig. 3). While it was anticipated that the profound hydrophobic character of the CF_3SF_4 group[9] could drive interfunctional group interactions in polar or aqueous media, the fluorous interactions[20] of those CF_3SF_4-containing side chains of TtsNVa suprisingly drove the conformation of the protected peptide **12** in $CDCl_3$. It is proposed that the head-to-tail alignment of the sidechains is promoted by CF_3SF_4 group induced dipolar interactions, an effect that would be maximized in the relatively non-polar solvent system. This conformational control is in marked contrast to the earlier experiments with SF_5-containing heptapeptides[12] where $CDCl_3$ solvation led to a loss of secondary structure.

Following deprotection, the low-energy structure of **5** was determined in DMSO-d_6 using the same NMR methods employed with **12.** A total of **235** manually and CANDID-assigned NOE cross-peaks were used in structural determination (See Supplementary information S2-4-2c, pS24 for **5**). The TtsNVa side chains of **5** remained in close proximity but had a slightly more distant head-to-head relationship as opposed to the head-to-tail organization of **12.** The reduction in the number of detectable NOE cross peaks was consistent with **5** assuming a less ordered conformation than **12**. Apparently deprotection and the more polar environment mitigated the intramolecular propensity of the CF_3SF_4 groups to associate.

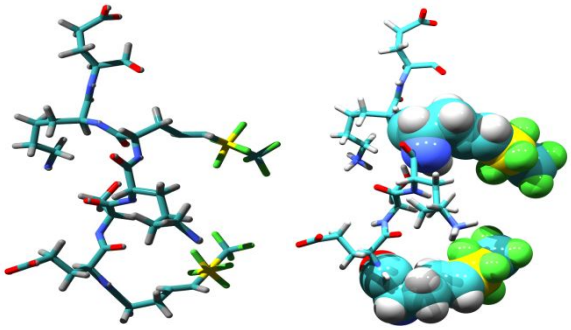


Figure 4. a. The NMR derived structure of the protected peptide **5** in DMSO-d_6. b. The lowest energy conformer of the best 20 structures is depicted with the atoms of the TtsNVa side chain represented as van der Waals spheres.

In summary, with only seven amino acid residues, the intramolecular influence of the CF_3SF_4 group is insufficient to provide the necessary driving force for a well-defined helical structure. Nonetheless the CF_3SF_4-containing side chains of TtsNVa effectively associate in both polar aprotic and non-polar solvent. Surprisingly the seven fluorines of the CF_3SF_4 group are sufficient to drive that organization. Selective incorporation of TtsNVa into longer peptides would be predicted to influence both the quaternary and tertiary structure of the peptides, independent of whether the peptides were in a highly polar aqueous environment or not. Along with the steric, electronic and dipolar effects, the remarkably potent combination of lipophobic and hydrophobic interactions that accompany CF_3SF_4 introduction can dramatically modify the properties of those substances bearing this unique group.

Conflicts of interest

There are no conflicts to declare.

Notes and references

1. (a) A. Niemz and D. A. Tirrell, *J. Am. Chem. Soc.*, 2001, **123**, 7407-7413; (b) B. Bilgicer, X. Xing and K. Kumar, *J. Am. Chem. Soc.*, 2001, **123**, 11815-11816; (c) B. Bilgicer and K. Kumar, *Tetrahedron*, 2002, **58**, 4105-4112.
2. (a) H.-Y. Lee, K.-H. Lee, H. M. Al-Hashimi and E. N. G. Marsh, *J. Am. Chem. Soc.*, 2006, **128**, 337-343; (b) M. Salwiczek and B. Koksch, *Chembiochem*, 2009, **10**, 2867-2870; (c) B. C. Buer, B. J. Levin and E. N. G. Marsh, *J. Am. Chem. Soc.*, 2012, **134**, 13027-13034; (d) B. C. Buer, J. L. Meagher, J. A. Stuckey and E. N. G. Marsh, *Proc. Natl. Acad. Sci. USA*, 2012, **109**, 4810-4815.
3. (a) H. Yang, C. Tian, D. Qiu, H. Tian, G. An and G. Li, *Org. Chem. Front.*, 2019, **6**, 2365-2370; (b) L. Zhu, J. Xiong, J. An, N. Chen, J. Xue and X. Jiang, *Org.Biomol. Chem.*, 2019, **17**, 3797-3804; (c) Y.-G. Lou, A.-J. Wang, L. Zhao, L.-F. He, X.-F. Li, C.-Y. He and X. Zhang, *Chem. Commun.*, 2019, **55**, 3705-3708; (d) C. Odar, M. Winkler and B. Wiltschi, *Biotechnol. J.*, 2015, **10**, 427-446; (e) E. N. G. Marsh, *Acc.*

Chem. Res., 2014, **47**, 2878-2886; (f) M. Salwiczek, E. Nyakatura, U. I. M. Gerling, S. Ye and B. Koksch, *Chem. Soc. Rev.*, 2012, **41**, 2135-2171; (g) U. I. M. Gerling, M. Salwiczek, C. D. Cadicamo, H. Erdbrink, C. Czekelius, S. L. Grage, P. Wadhwani, A. S. Ulrich, M. Behrends, G. Haufe and B. Koksch, *Chem. Sci.*, 2014, **5**, 819-830; (h) A. A. Berger, J.-S. Voeller, N. Budisa and B. Koksch, *Acc. Chem. Res.*, 2017, **50**, 2093-2103.

4. (a) M. A. Molski, J. L. Goodman, C. J. Craig, H. Meng, K. Kumar and A. Schepartz, *J. Am. Chem. Soc.*, 2010, **132**, 3658-3659; (b) L. Merkel and N. Budisa, *Org. Biomol. Chem.*, 2012, **10**, 7241-7261; (c) E. K. Nyakatura, O. Reimann, T. Vagt, M. Salwiczek and B. Koksch, *RSC Adv.*, 2013, **3**, 6319-6322; (d) C. Jäckel, M. Salwiczek and B. Koksch, *Angew. Chem. Int. Ed.*, 2006, **45**, 4198-4203.
5. H.-P. Chiu, Y. Suzuki, D. Gullickson, R. Ahmad, B. Kokona, R. Fairman and R. P. Cheng, *J. Am. Chem. Soc.*, 2006, **128**, 15556-15557.
6. (a) S. S. Pendley, Y. B. Yu and T. E. Cheatham, 3rd, *Proteins*, 2009, **74**, 612-629; (b) C. Jaeckel, W. Seufert, S. Thust and B. Koksch, *Chembiochem*, 2004, **5**, 717-720.
7. (a) S. Son, I. C. Tanrikulu and D. A. Tirrell, *Chembiochem*, 2006, **7**, 1251-1257; (b) B. C. Buer, R. de la Salud-Bea, H. M. Al Hashimi and E. N. G. Marsh, *Biochemistry*, 2009, **48**, 10810-10817; (c) J. K. Montclare, S. Son, G. A. Clark, K. Kumar and D. A. Tirrell, *Chembiochem*, 2009, **10**, 84-86.
8. A. Ikeda, L. Zhong, P. R. Savoie, C. N. von Hahmann, W. Zheng and J. T. Welch, *Eur. J. Org. Chem.*, 2018, **2018**, 772-780.
9. P. Kirsch and A. Hahn, *Eur. J. Org. Chem.*, 2006, **2006**, 1125-1131.
10. (a) K. A. Dill, *Biochemistry*, 1990, **29**, 7133-7155; (b) P. L. Privalov and S. J. Gill, *Advances in Protein Chemistry*, 1988, **39**, 191-234; (c) J. Gao, S. Qiao and G. M. Whitesides, *J. Med. Chem.*, 1995, **38**, 2292-2301; (d) L. R. Pratt, *Annu. Rev. Phys. Chem.*, 1985, **36**, 433-449; (e) A. Ben-Naim, *Hydrophobic Interactions*, Plenum Press, 1980; (f) C. Tanford, *The Hydrophobic Effect: Formation of Micelles and Biological Membranes 2d Ed*, J. Wiley., 1980.
11. J. C. Biffinger, H. W. Kim and S. G. DiMagno, *Chembiochem*, 2004, **5**, 622-627.
12. D. S. Lim, J.-H. Lin and J. T. Welch, *Eur. J. Org. Chem.*, 2012, **2012**, 3946-3954,.
13. (a) P. Güntert, ed., *Automated NMR Structure Calculation With CYANA*, 2004; (b) P. Guntert, in *Protein NMR Spectroscopy: Practical Techniques and Applications*, John Wiley & Sons, Ltd, 2011, pp. 159-192.
14. (a) J. T. Welch, in *Fluorine in Pharmaceutical and Medicinal Chemistry: From Biophysical Aspects to Clinical Applications* eds. V. Gouverneur and K. Muellller, Imperial College Press Co. , 2012, vol. 5, pp. 171-204; (b) C. N. von Hahmann, P. R. Savoie and J. T. Welch, *Curr. Org. Chem.*, 2015, **19**, 1592-1618; (c) P. R. Savoie and J. T. Welch, *Chem. Rev. (Washington, DC, U. S.)*, 2015, **115**, 1130-1190; (d) S. Altomonte and M. Zanda, *J. Fluorine Chem.*, 2012, **143**, 57-93.
15. T. R. Hoye, C. S. Jeffrey and F. Shao, *Nature Protocols*, 2007, **2**, 2451.
16. J. Cavanagh, W. Fairbrother, I. I. I. A. G. Palmer, N. Skelton and Editors, *Protein NMR Spectroscopy: Principles and Practice*, Academic, San Diego, 1995.
17. (a) P. Güntert, in *BioNMR in Drug Research*, Wiley-VCH Verlag GmbH & Co. KGaA, 2003, pp. 39-66; (b) P. Guntert, *Prog. Nucl. Mag. Res. Sp.*, 2003, **43**, 105-125; (c) C. Mumenthaler, P. Güntert, W. Braun and K. Wüthrich, *J. Biomol. NMR*, 1997, **10**, 351-362.
18. T. Ikeya, J.-G. Jee, Y. Shigemitsu, J. Hamatsu, M. Mishima, Y. Ito, M. Kainosho and P. Güntert, *J. Biomol. NMR*, 2011, **50**, 137-146.
19. G. Cornilescu, F. Delaglio and A. Bax, *J. Biomol. NMR*, 1999, **13**, 289-302.
20. (a) M. Cametti, B. Crousse, P. Metrangolo, R. Milani and G. Resnati, *Chem. Soc. Rev.*, 2012, **41**, 31-42; (b) D. P. Curran, in *Handbook of Fluorous Chemistry*, Wiley-VCH Verlag GmbH & Co. KGaA, 2004, pp. 128-155; (c) I. T. Horvath, D. P. Curran and J. A. Gladysz, in *Handbook of Fluorous Chemistry*, Wiley-VCH Verlag GmbH & Co. KGaA, 2004, pp. 1-4; (d) K. Olofsson and M. Larhed, in *Handbook of Fluorous Chemistry*, Wiley-VCH Verlag GmbH & Co. KGaA, 2004, pp. 359-365; (e) J.-M. Vincent, *Chem. Commun. (Cambridge, U. K.)*, 2012, **48**, 11382-11391; (f) W. Zhang, in *Current Fluoroorganic Chemistry, ACS Symp. Ser.*, American Chemical Society, 2007, vol. 949, pp. 207-220.