Open access • Journal Article • DOI:10.1039/C1CS15241F

Fluorinated amino acids: compatibility with native protein structures and effects on protein–protein interactions — Source link

Mario Salwiczek, Elisabeth K. Nyakatura, Ulla I. M. Gerling, Shijie Ye ...+1 more authors

Institutions: Free University of Berlin

Published on: 27 Feb 2012 - Chemical Society Reviews (The Royal Society of Chemistry)

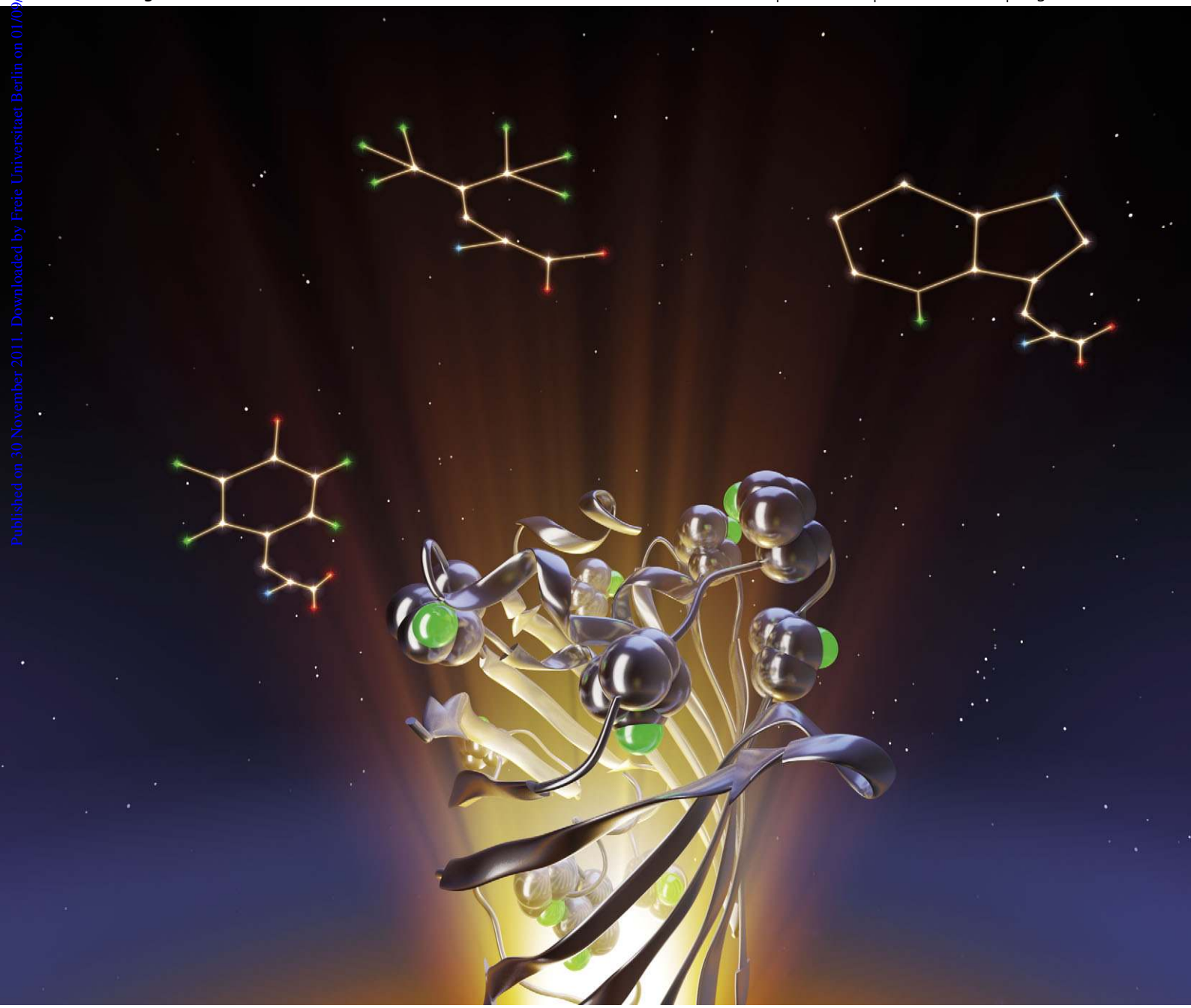
Topics: Protein structure and Amino acid

Chem Soc Rev

Chemical Society Reviews

www.rsc.org/chemsocrev

Volume 41 | Number 6 | 21 March 2012 | Pages 1989–2520



ISSN 0306-0012

CRITICAL REVIEW

Mario Salwiczek, Elisabeth K. Nyakatura, Ulla I. M. Gerling, Shijie Ye and Beate Koksch

Fluorinated amino acids: compatibility with native protein structures and effects on protein–protein interactions

RSC Publishing

0306-0012(2012)41:6;1-R

Chem Soc Rev

Cite this: *Chem. Soc. Rev.*, 2012, **41**, 2135–2171

www.rsc.org/csr

CRITICAL REVIEW

Fluorinated amino acids: compatibility with native protein structures and effects on protein–protein interactions

Mario Salwiczek,* Elisabeth K. Nyakatura, Ulla I. M. Gerling, Shijie Ye and Beate Koksch*

Received 6th September 2011
DOI: 10.1039/c1cs15241f

Fluorinated analogues of the canonical α-L-amino acids have gained widespread attention as building blocks that may endow peptides and proteins with advantageous biophysical, chemical and biological properties. This *critical review* covers the literature dealing with investigations of peptides and proteins containing fluorinated analogues of the canonical amino acids published over the course of the past decade including the late nineties. It focuses on side-chain fluorinated amino acids, the carbon backbone of which is identical to their natural analogues. Each class of amino acids—aliphatic, aromatic, charged and polar as well as proline—is presented in a separate section. General effects of fluorine on essential properties such as hydrophobicity, acidity/basicity and conformation of the specific side chains and the impact of these altered properties on stability, folding kinetics and activity of peptides and proteins are discussed (245 references).

1. Introduction

Fluorine has emerged as a "magic element" in medicinal chemistry as well as in crop and materials science. Fluorine substitutions are now considered a standard strategy for modulating the properties of chemical leads. In many cases this strategy is so successful that more than 20% of pharmaceuticals nowadays contain at least one fluorine atom.[1] The great success of fluorine substitutions as an application is based on its unique stereoelectronic properties that arise from an unprecedented combination of small size, very low polarizability and the strongest inductive effect found among the chemical elements. Although most of its effects as an organic substituent are expectedly context-dependent, the presence of fluorine within a molecule often favourably and in many cases predictably alters biophysical and chemical properties such as hydrophobicity, acidity/basicity, reactivity and conformation. Moreover, perfluorinated molecules exhibit a tendency to

Department of Biology, Chemistry, Pharmacy, Institute of Chemistry and Biochemistry, Freie Universität Berlin, Takustr. 3, 14195 Berlin, Germany. E-mail: Beate.Koksch@fu-berlin.de, m.salwiczek@fu-berlin.de

Mario Salwiczek

Mario Salwiczek studied chemistry with a scholarship of the "Fonds der chemischen Industrie" at the University of Leipzig. In 2010 he received a PhD in bioorganic chemistry from Freie Universität Berlin where he was working with Prof. Koksch. His work focused on the consequences of single fluoroamino acid substitutions on the thermodynamics and kinetics of coiled-coil interactions. He is currently a postdoctoral fellow at the Center for Supramolecular Interactions. Mario's research interest is to understand and tune the biological activity of peptides that bear potential both as pharmaceutically active agents and biocompatible materials.

Elisabeth K. Nyakatura

Elisabeth K. Nyakatura received her BSc in chemistry from the Freie Universität Berlin, Germany, in 2007. She then joined the London School of Hygiene and Tropical Medicine and successfully completed a MSc programme in Molecular Biology of Infectious Diseases. After graduating from London in 2008, she returned to the Freie Universität to pursue her PhD in the group of Prof. Koksch. Her research is now focused on using phage display to identify the preferred interaction partners of side chain fluorinated amino acids in a native protein environment. It is this interface of chemistry and molecular biology that fuels Elisabeth's interest in using chemical means to address pathogen-host interactions.

segregate from both, aqueous and hydrophobic environments affording a third, the "fluorous" phase that has proven broad applicability in the fields of catalysis and separation processes. It has been anticipated that fluorine's effects would also be of use to beneficially modulate the properties of peptides and proteins. Peptides have shown great potential as highly active pharmaceuticals. Considering possible industrial applications as catalysts or biocompatible materials, enhancing their stability and bioavailability are major goals in protein engineering. The combined efforts of solid phase peptide synthesis, chemoselective peptide ligation,[2] non-natural protein expression[3] as well as protein semi-synthesis[4] enable site-specific as well as global incorporation of non-natural amino acids into peptides and proteins. However, the intellectual and practical challenge of non-natural peptide and protein synthesis notwithstanding, the synthetic accessibility of fluorinated peptides and proteins is mostly limited by the availability of appropriate building blocks for solid phase peptide synthesis or their compatibility with the protein expression machinery. Although the synthetic strategies that provide access to various fluorinated analogues of the canonical amino acids are well established,[5,6] most studies focus on the application of aliphatic and aromatic fluorinated amino acids. Moreover, fluorinated proline analogues have proven their applicability to study structural aspects of polyprolines. Analogues of trifunctional amino acids, such as polar and charged structures, have not gained as much attention mainly due to the chemical instability of some of these analogues as well as difficulties in their synthesis.

The aim of this *critical review* is to provide an overview of the achievements in the field of fluorinated amino acid research with a focus on the following aspects: (1) side-chain fluorinated α-L-amino acids that have been used to modify peptides and proteins; (2) effects of fluorine substitutions on the intrinsic properties of the amino acids; (3) peptides and proteins that have been modified and (4) how the altered properties of the fluorinated side chains translate into properties of the modified structures. We focus on side-chain fluorinated analogues of the canonical L-amino acids, *i.e.* amino acids that carry the same number of carbon atoms as the natural analogues that are encoded within the genome. The synthesis of C^{α}-fluoroalkyl amino acids[7] and their effects on peptide/protein interactions[8] have recently been reviewed by our group and, therefore, are not part of the present overview. Moreover, studies dealing with fluorinated amino acids as ^{19}F-NMR probes[9,10] that do not explicitly address the chemical/biophysical effects of fluorine substituents are not included. Each class of amino acids—aliphatic/methionine, aromatic, charged and polar as well as proline—is presented in a separate section. The consequences of fluorine substitutions in organic molecules have been the issue of several very recent reviews.[11,12] However, those properties that are most important to realize in order to understand their effects on protein folding and stability are briefly discussed at the beginning of each section.

A consistent three-letter code for non-natural amino acids does not exist and, often, different abbreviations are used by different authors. We propose to use abbreviations such as those used throughout this review that clearly indicate the position and number of fluorine substitutions as well the stereochemistry at the substituted carbon atoms (where applicable).

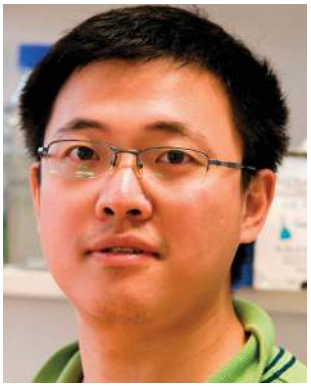
2. Analogues of aliphatic amino acids and methionine

2.1 General properties

Aliphatic amino acids (Ala, Val, Ile and Leu) carry chemically inert side chains that usually accumulate at certain sites within

Ulla I. M. Gerling

Ulla I. M. Gerling was born in 1982 in Germany. She studied chemistry at Freie Universität Berlin and graduated in 2009 with an MSc in chemistry working with Prof. Koksch. She is currently a PhD student in the same group. Ulla investigates the properties and morphology of fluorinated amyloid forming peptides.



Shijie Ye

Shijie Ye was born in 1980 in Shanghai, China. From 1999 to 2001 he studied human medicine at Shanghai Second Medical University (now School of Medicine, Shanghai JiaoTong University). From 2003 to 2007 he studied biochemistry at Freie Universität Berlin. He completed his diploma thesis "Cell-free expression system for membrane receptors" with Prof. Holger Stark at Goethe-Universität Frankfurt am Main and received his diploma in 2007. Shijie joined Prof. Koksch's group in 2008 to work on his PhD thesis on site-specific incorporation of fluorinated amino acids into proteins.

Beate Koksch

Beate Koksch received her diploma in Chemistry and her PhD in biochemistry from the University of Leipzig. She was a DFG research fellow in the laboratories of Prof. M. R. Ghadiri and Prof. C. F. Barbas, III at TSRI, La Jolla, California. In 2000 she returned to the University of Leipzig to start her independent career under the mentorship of Prof. Dr K. Burger. In 2004 she assumed a position as Professor for Organic and Natural Product Chemistry at the Freie Universität Berlin. Her research focuses on the development of peptide models to study peptide and protein folding.

the native structure of proteins where they form a hydrophobic core. Hydrophobic interactions stabilize the native state as well as substrate binding in addition to so-called packing interactions and provide a large portion of the entropic component of the driving force for folding/binding by the hydrophobic effect. It is a frequently produced argument that hydrogen-to-fluorine substitutions, while often discussed to be conservative in terms of steric size, provide the means for increasing hydrophobicity of aliphatic hydrocarbons and thus may be used to stabilize the protein structure. However, due to the high electronegativity of the fluorine atom, one major distinction of the C–F and the C–H bond is that the C–F bond exhibits a considerable dipole moment which is reversed in orientation compared to the small C–H dipole. Single aliphatic fluorine substitutions actually reduce rather than increase hydrophobicity, and a critical number of fluorine atoms per alkyl chain is required to achieve what could be called "hyper-hydrophobicity".[13] One plausible reason for the extreme hydrophobicity of highly fluorinated aliphatic hydrocarbons is fluorine's inherently low polarizability that makes it reluctant to dispersive interactions with water *and* other hydrocarbons. Therefore, perfluorocarbons tend to segregate from both hydrophilic *and* lipophilic environments. The exceptional phase segregation behaviour of perfluoroalkyl groups offers interesting prospects in protein engineering (*vide infra*). Despite its strong dipole, the C–F bond hardly ever acts as a hydrogen bond acceptor. However, it has been found that orthogonal polar interactions with carbonyl groups frequently occur in small molecule–protein co-crystals.[11]

Most fluorinated aliphatic amino acids that have been incorporated into peptides or proteins have one or two CH_3 groups replaced by CF_3 groups. Although hydrogen and fluorine are often discussed to be almost isosteric, an extrapolation to higher fluorinated alkyl groups would be misleading. In fact, the CF_3 group approximates twice the van der Waals volume of a CH_3 group and the steric effects of the CF_3 group were shown to be close to that of an isopropyl group[14] or to larger substituents such as *sec*-butyl and cyclohexyl groups.[15] Size and shape of a side chain determine its ability to engage in packing interactions with other side chains. Moreover, size and hydrophobicity closely correlate. Using analogues of α-L-aminobutyric acid Samsonov *et al.* have shown that the impact of an increasing fluorine content on the hydrophobicity of aliphatic amino acids is indeed larger than a mere increase in steric size by elongation and branching of the side chain.[13] Hydrophobicity is frequently evaluated on the basis of distribution coefficients between water and organic solvents such as octanol or heptanol. However, while consistent hydrophobicity scales based on distribution coefficients are available for the canonical amino acids, only few fluorinated analogues have been investigated. Therefore, a complete hydrophobicity scale for the fluorinated analogues of aliphatic amino acids is not available at this time.

Because fluorine exerts a strong inductive and hyperconjugative effect, it influences the pK_a values of nearby functional groups.[16,17] In general, the presence of fluorine increases acidity and its effect is stronger on the α-amino group than on the carboxyl group. In line with these alterations, the strengths of hydrogen bonds that are crucial to secondary structure formation are affected. However, the individual impacts of hydrogen bond strength, conformation and hydrophobicity

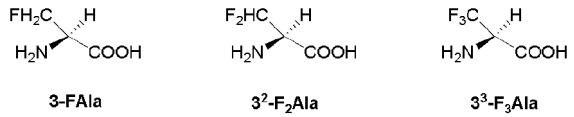


Fig. 1 Structures of side-chain fluorinated L-alanine analogues.

to structure formation and stability are hard to disentangle. Thus, although the impact on individual properties of aliphatic amino acids such as hydrophobicity and pK_a may be assessable, the combined effects on protein structure and stability are not easily predicted.

2.2 Alanine

Although the fluorinated analogues of alanine (Fig. 1) have been described,[5] applications in peptide and protein science are rare. While the free analogues are stable, the activated species tend to eliminate hydrogen fluoride leading to formation of dehydroalanine under basic conditions. Moreover, the very close proximity of the fluorine atoms to the N^{α} reduces its nucleophilicity to an extent that it seriously affects the reactivity in peptide coupling reactions.

There are only a few examples of 3-FAla and 3^3-F_3Ala containing small peptides found in the literature. Mitra *et al.* described the synthesis of tetra-L-alanine analogues containing 3^3-F_3Ala at positions 1, 2 and 3.[18] They tested the analogue modified at position 2 for its binding to an antipoly-L-alanine antibody and found no difference compared to the parental Ala_4 sequence. The other analogues modified at position 1 and 3 have not been investigated yet.

The chemical instability of fluoroalanines may be somewhat unfortunate for peptide synthesis. However, this instability may be used as a synthetic tool. For example, the D-enantiomer of 3-FAla has been used to produce cyclosporin A (CyA) analogues with modifications at position 8.[19,20] Following dehydrofluorination the resulting double bond was amenable to site and stereospecific modifications using different electrophiles such as thiols for Michael reactions. Importantly, the fluorinated analogue itself showed 80% of the native suppressive activity which makes it one of the most active analogues of CyA.

The N-carboxy anhydride of 3^3-F_3Ala has been investigated for its polymerization properties by Dessipri and Tirrell.[21] The authors generated homopolymers of D,L-3^3-F_3Ala as well as copolymers with glutamate and assessed the surface energy of these polypeptides. As expected for highly fluorinated molecules, they found very low surface energies for these polymers as shown by the very large contact angles of water on thin films of the polymers.

Nevertheless, in contrast to other aliphatic amino acids, fluorinated analogues of alanine have not directly been investigated for their effect on protein stability and structure. Recent and ongoing endeavours to facilitate their synthesis and incorporation into peptides[22] will certainly motivate further developments. Their intrinsic instability, however, makes them less attractive as candidates for protein structure modifications.

2.3 Leucine, isoleucine and valine

Fluorinated analogues of isoleucine, leucine and valine (Fig. 2) are the most commonly used building blocks for peptide and

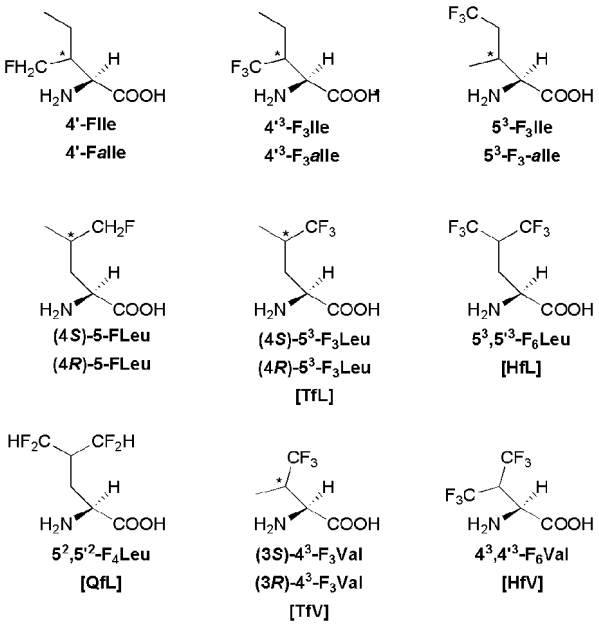


Fig. 2 Side chain fluorinated analogues of the aliphatic L-amino acids: Ile, Leu and Val that have been incorporated into peptides and proteins (commonly used abbreviations are given in brackets, *a* = *allo*).

protein modifications. Their syntheses are mostly straightforward and major difficulties in chemical peptide synthesis as well as protein expression are rarely observed. The incorporation of a heterosubstituent such as fluorine may generate new stereocentres within the side chain offering the opportunity to study the effects of stereochemistry on the protein structure and stability.

2.3.1 Helical structures

Membrane active peptides. Membrane active peptides/antimicrobial peptides (AMPs) provide a starting point for the development of highly active and specific antimicrobial agents. The structural basis for their ability to interact with cell membranes is their amphiphilic, often helical structure with spatially separated hydrophobic and hydrophilic residues. A substitution of hydrophobic residues by fluorinated analogues has been employed by several groups as a strategy to increase their activity and stability.

Gottler *et al.* globally replaced the Leu and Ile residues of pexiganan (H_2N-**G**I**G**K**F**L**K**K**A**K**K**F**G**K**A**F**V**K**I**L**K**K-COOH),[23] a synthetic analogue of magainin-2, with 5^3,$5'^3$-F_6Leu. The peptide retained its antimicrobial activity while remaining haemolytically inactive. Furthermore, the fluorinated analogue showed improved stability towards trypsin as well as chymotrypsin digestion in the presence of a membrane environment. In the absence of a membrane, however, the peptide was as rapidly degraded as the all-native wild type. Thus, the stabilizing effect was attributed to an increased tendency of the peptide to self-assemble within the lipid bilayer due to fluorination.

Niemz and Tirrell investigated several fluorinated variants of the bee venom melittin (H_2N-G**I**G**A**V**L**K**V**L**T**T**G**L**P**A**L**I**S**W**I**K-RKRQQ-$CONH_2$) for their self-association and membrane-binding behaviour.[24] Several analogues of the peptide were synthesized in which either position 9 (M-9TfL) or position 13 (M-13TfL) was replaced by a diastereomeric mixture of 5^3-F_3Leu. In addition, a tetrasubstituted variant with 5^3-F_3Leu at all four positions (6, 9, 13 and 16, M-allTfL) was prepared. The authors attempted a separation of the diastereomeric peptide mixtures which was only successful for M-13TfL, while the M-9TfL diastereomers were inseparable. Moreover, the 16 possible diastereomers of M-allTfL merely eluted as two fractions. Thus, most results remained inconclusive as to the impact of stereochemistry. In general, all peptides showed an increased membrane affinity that was shown to be the result of enhanced self-association of the peptides. The authors argued that the more hydrophobic character of the fluorinated peptides would result in an increased partition coefficient without altering the apparent effective charge of the membrane-bound peptide. Instead a marked decrease in the apparent effective charge was found. Therefore, the enhanced self-association is driven by the more hydrophobic and at the same time lipophobic character of the highly fluorinated alkyl side chains that, apparently, disfavour the lipophilic membrane environment. The results for the two diastereomers of M-13TfL were shown to be different. However, since high-resolution structures were not available, a conclusive interpretation of the differences has yet to be provided. The conclusion, that fluorination of aliphatic residues may induce lipophobicity in line with increasing hydrophobicity was already drawn earlier by Arai *et al.* who tested a variant of gramicidin S (GS, *cyclo*[**Val**OrnLeuDPhePro]$_2$) in which the two Val residues were replaced by 4^3,$4'^3$-F_6Val.[25] Although the typical activity profile as well as the structure of GS was not altered upon fluorination, the analogue displayed an eight- to sixteen-fold lower activity. This finding was interpreted as a weaker peptide–membrane interaction caused by the lipophobicity of the highly fluorinated Val analogues.

In another comparative study employing two AMPs with different membrane binding mechanisms, buforin-2 (H_2N-TRSSRAGLQFPVGRVHR**LL**RK-COOH) and magainin-2 (H_2N-GIGKF**L**HA**A**KKF**A**KAF**V**AE**I**MNS-$CONH_2$), Meng and Kumar showed that the impacts of fluorination on stability and activity may not easily be extrapolated from one model to another.[26] Several variants of both peptides containing different levels of 5^3,$5'^3$-F_6Leu as a replacement for Leu, Ile and Val (as well as Ala in one case) at the nonpolar peptide–membrane interface were prepared. All fluorinated variants except the globally fluorinated magainin-2 analogue showed similar or enhanced antimicrobial activity compared to their native analogues. The reduction in activity of the fully fluorinated magainin-2 was explained by its markedly enhanced self-assembly in aqueous solution which presumably hampers its initial interaction with the membrane. In contrast to magainin-2 analogues, the activity of the buforin-2 series was shown to correlate with hydrophobicity. Moreover, while all buforin-2 analogues were haemolytically inactive, fluorination induced higher haemolytic activity in the case of magainin-2. The beneficial effect of fluorine was also shown by the fact that truncated buforin-2 variants, being usually inactive, regain activity upon fluorination. Interestingly and in contrast to the findings for the pexiganan analogues described by Gottler *et al.*,[23] the peptides in this study exhibited increased proteolytic stability even in the absence of a membrane. In addition, different cleavage patterns were observed for some of the fluorinated variants.

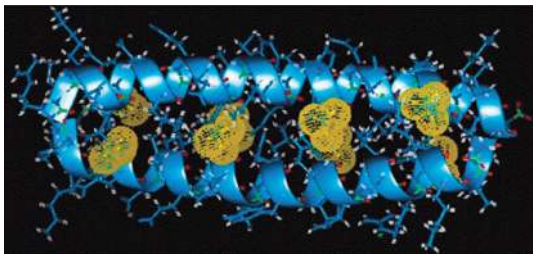


Fig. 3 Modelled structure of GCN4 with trifluoroleucine at the four d positions. Yellow spheres represent the van der Waals radii of the fluorine atoms (reproduced with permission from Tang *et al.*[28] Copyright 2011 American Chemical Society).

Coiled-coil peptides/helical bundles. Coiled-coil peptides have gained widespread attention as model systems for studying the impact of fluorination on protein–protein interactions. Coiled-coil structures are based on a (pseudo-) repetitive sequence $(abcdefg)_n$, the so-called heptad repeat, in which hydrophobic side chains are primarily located at the a- and d-positions on one side of the helix, while most of the other positions are hydrophilic or charged. By applying this general principle, self-assembling helical bundles of different oligomerization states and specificities can be designed *de novo*.[27]

Tang *et al.* incorporated diastereomeric mixtures of 5^3-F_3Leu into all four d positions of the dimeric C-terminal coiled-coil subdomain of GCN4-p1, a naturally occurring transcription factor by means of solid phase peptide synthesis.[28] Furthermore they produced a fluorinated analogue of the full-length transcription factor for DNA binding studies. Due to the parallel alignment of the two helices the fluorinated side chains in d-positions pack against the d′-positions of the adjacent helix so that the fluorinated side chains directly face each other (Fig. 3). The fluorinated peptides showed enhanced thermal stability and a higher resistance towards denaturation by chaotropic agents. Importantly, fluorination did not hamper DNA-binding affinity.

Son *et al.* studied the impact of Ile and Val substitutions by 5^3-F_3Ile and (3*R*)-4^3-F_3Val, respectively, this time at the a-position of a slightly adapted analogue of GCN4.[29] The respective variants were successfully expressed in *E. coli*. It had been observed earlier that the (3*S*)-isomer of 4^3-F_3Val does not support protein biosynthesis in this bacterial host[30] and, thus, only (3*R*)-4^3-F_3Val was included in these studies. Both fluorinated peptides were shown to exhibit DNA-binding affinities and specificities similar to the all-native sequence. As observed before, fluorination increased the intrinsic stability of the dimer, albeit to a markedly different extent, when Ile and Val analogues were compared. The replacement of the δ-CH_3 group of Ile by CF_3 showed an approximately eight-fold higher stabilization than the replacement of the γ-CH_3 of Val. This finding could be explained with the substantial loss of side-chain entropy of Val due to steric clashes between the significantly larger γ-CF_3 group and the helix backbone which does not occur in the case of Ile.

Diastereomeric 5^3-F_3Leu and 5^3,$5'^3$-F_6Leu were also incorporated into the d-positions of an artificial coiled-coil dimer by *in vivo* procedures.[31,32] 5^3-F_3Leu could be incorporated into the recombinant protein employing a Leu-auxotrophic *E. coli* strain in good yields, while 5^3,$5'^3$-F_6Leu did not initially support biosynthesis. The authors applied an engineered bacterial host that overexpressed leucyl-tRNA-synthetase to overcome the unfavourable activation kinetics for 5^3,$5'^3$-F_6Leu.[32] The all-hexafluoroleucine analogue was even more stable than the all-trifluoroleucine variant, indicating that stability and fluorine content closely correlate. However, in most of these studies a mixture of (4*S*)- and (4*R*)- of 5^3-F_3Leu was used which raises the question to what extent the pure isomers would differ in their stabilizing effect. Thus, the impact of 5^3-F_3Leu stereochemistry on coiled-coil stability was recently investigated by Montclare *et al.*[33] To this end, the earlier described artificial coiled coil was used as a model into which both diastereomers were separately incorporated at all d-positions by means of protein expression. The 4*S* isomer was activated at slightly higher rates by leucyl-tRNA-synthetase during expression in *E. coli*. However, both isomeric peptides were produced in good yields. It was found that the extent to which both diastereomers stabilize the dimer is practically identical as judged by their thermal melting points. An equimolar mixture of the peptides containing either the 4*S*- or the 4*R*-isomer of 5^3-F_3Leu was shown to exhibit even higher thermal stability. Importantly, the melting point of the peptide mixture was identical to the earlier investigated dimer that contained a diastereomeric mixture of 5^3-F_3Leu. These results, however, remained without further explanation, probably because high-resolution structures that could potentially reveal the packing patterns of the fluorinated side chains within the homo- and heterodimers were not available. Moreover, the investigations provided no insight as to which species were present within the equimolar mixture of the isomeric peptides. An increased melting temperature could indicate the presence of a new species. However, its population at equilibrium and its identity was not unequivocally confirmed. Another important point to make is that protein expression employing auxotrophic bacterial strains does usually not yield uniform products. In the studies described above the level of non-natural amino acid incorporation rarely exceeded 90%. Since the products were not purified, all data reflect the properties of a product mixture.

A fully fluorinated analogue of the GCN4-p1 coiled-coil domain was presented by Bilgiçer *et al.*[34] In that case all a-positions and all d-positions were endowed with diastereomeric mixtures of 4^3-F_3Val and 5^3-F_3Leu, respectively, in parallel. Since the sequences of GCN4 analogues are not 100% identical in all these studies direct comparisons should be made carefully. However, the thermal stabilization imparted by substituting a- and d-positions at the same time does not exceed the effects of the individual 5^3-F_3Leu substitution at the d-position that was observed earlier. A comparison of thermodynamic stabilities in terms of free energy would certainly be more significant. However, as standard conditions for reporting such data (temperature, buffer, ion strength, pH) have not been defined yet, the absolute values observed by different authors are not directly comparable. Also, the determination of standard thermodynamic values often requires large extrapolations and the extent to which such extrapolations are performed varies quite significantly.

The same authors extended their design to study the impact of fluorination on the specificity of coiled-coil interactions.[35,36]

They designed a dimeric model (HH) bridged by disulfide bonds that contained Leu at all a- and d-positions which were then replaced by $5^3,5'^3$-F_6Leu to afford FF. The stability of their constructs increased in the order HH < FH < FF. The disulfide bridge furthermore allowed for constructing heterodimers (FH). When these FH heterodimers were subjected to redox conditions, the authors observed preferential formation of the least stable HH and the most stable FF species and only trace amounts of FH. This segregation behaviour led to the conclusion that the fluorinated side chains are not only more hydrophobic but apparently disfavour interactions with hydrocarbon side chains. The findings were explained by invoking the "fluorous effect", the same property that is responsible for the immiscibility of perfluorinated and hydrocarbon solvents under standard conditions. A thermodynamic sedimentation equilibrium analysis, however, also showed that a complete fluorination of the hydrophobic core induced an oligomeric switch from a dimer to a tetramer. The same self-sorting behaviour has been observed in the presence of detergent micelles[37] and phospholipid bilayers.[38] Due to their highly increased hydrophobicity and lipophobicity fluorinated leucines segregate from the lipophilic environment and, therefore, tend to enhance helix–helix interactions in these membrane environments.

The conclusion that a specific "interaction" of fluorous phases may be invoked to explain the above findings has been challenged by the group of Neil Marsh. They employed a tetrameric, antiparallel coiled-coil system to investigate the effects of Leu to $5^3,5'^3$-F_6Leu substitutions on its stability and self-sorting behaviour.[39,40] Their sequence contained Leu in all a and d positions that were substituted in a stepwise fashion, layer by layer (Fig. 4). The stabilization of the two central layers (0.3 kcal mol^{-1} per residue) almost exactly matches the by 0.4 kcal mol^{-1} more favoured partitioning of $5^3,5'^3$-F_6Leu into organic solvents.[39] Thus, the authors concluded that the stabilizing effects were simply caused by the increased hydrophobicity of hexafluoroleucine rather than a specific "fluorine–fluorine interaction".

Further variants of the four-helix bundle in which four to six layers of the hydrophobic core were replaced (Fig. 4) have been investigated.[40] Interestingly, the introduction of an additional layer of $5^3,5'^3$-F_6Leu did not result in the same increase of stability as observed for two layers. In fact, only 0.12 kcal mol^{-1} per residue stabilization was observed. The fully fluorinated variant was furthermore shown to be more stable towards chymotrypsin and trypsin digestion. In a series of NMR experiments they examined the effect of fluorination on the conformational dynamics and observed that an increasing degree of fluorination resulted in a higher degree of order—the peptide assemblies became less dynamic. Thus, global replacement of all leucines by $5^3,5'^3$-F_6Leu results in overpacking of the core, which antagonizes the stabilizing effects of fluorine. This may also explain the oligomeric switch reported earlier by Bilgiçer *et al.*[35] To prove their hypothesis that fluorine's stabilizing effect may solely be ascribed to its effect on hydrophobicity the authors attempted to disrupt the helical bundles using different hydrocarbon and fluorocarbon solvents.[41] While the addition of 2-propanol or ethanol quickly unfolded the unmodified four-helix bundle leaving the fully fluorinated analogue unaffected, highly fluorinated solvents such as trifluoroethanol (TFE) disrupted both, however, and left behind monomeric helical peptides. They argued that if a specific fluorous effect existed, the fluorinated solvent would have preferentially disrupted the fluorinated peptide. A self-segregation of the fluorinated from the native peptide could not be observed either, since mixed forms were found by ^{19}F-NMR-experiments. However this interpretation is disputable as TFE has been shown to induce monomeric α-helical structures for several structurally distinct peptides.[42,43] It may be assumed that the mechanism by which TFE affects the structure of peptides is different from simple denaturation by other organic solvents. It is nevertheless interesting that the authors did not observe self-segregation as judged from NMR titrations of the fluorinated against the non-fluorinated peptide. One major distinction of their coiled-coil model is that the helices are aligned in an antiparallel rather than a parallel manner and it is a well-known fact that the orientation of helices within a coiled-coil assembly determines the packing of the side chains within the hydrophobic core.

The formation of mixed species may be explained by a more favourable packing of the fluorinated against the non-fluorinated side chains that possibly outweigh the fluorous effect. Certainly, it does not unequivocally rule out the existence of the fluorous effect. A direct support of the above hypothesis, however, may be inferred form a recent publication by Buer *et al.*[44] The authors studied variants of the same four-helix bundle in which either the a or d positions were fully occupied with $5^3,5'^3$-F_6Leu thereby creating bundles in which Leu and $5^3,5'^3$-F_6Leu alternately pack against each other (Fig. 5). They found that the stabilization per hexafluoroleucine residue for these peptides is larger than in the fully fluorinated variant and attributed this finding to a more favourable packing. The fluorous effect may be present

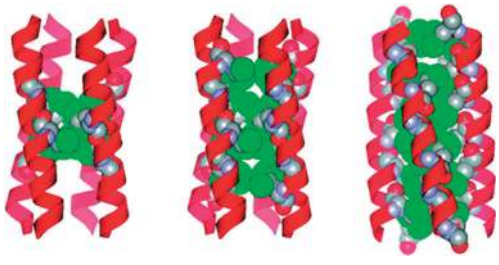


Fig. 4 Models of the four-helix bundle containing $5^3,5'^3$-F_6Leu in two, four and six layers of the hydrophobic core (adapted with permission from Lee *et al.*[40] Copyright 2011, American Chemical Society).

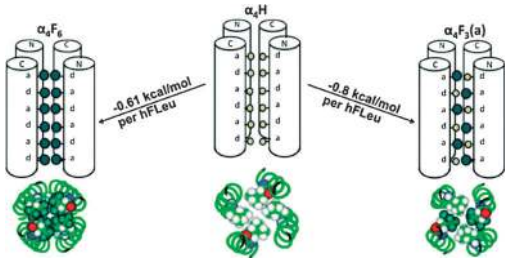


Fig. 5 Schematic representation of the tetrameric helical bundle showing the different substitution patterns and the per $5^3,5'^3$-F_6Leu residue stabilization involved (reproduced with permission from Buer *et al.*[44] Copyright 2011, American Chemical Society).

but packing interactions make large contributions to the stabilization of the protein structure that are not to be underestimated.

Pendley *et al.* recently presented a theoretical study of a highly fluorinated parallel coiled-coil heterodimer that contained 5^3,$5'^3$-F_6Leu at all d-positions of both monomers.[45] Their simulations confirm the experimental findings that fluorination of Leu increases stability of the helical interactions thereby enhancing coiled-coil formation. Interestingly, their studies suggest that the dimer is not only stabilized by the increased hydrophobicity of the fluorinated side chain. They also found that the orientation of 5^3,$5'^3$-F_6Leu may allow for enhanced favourable electrostatic interactions between charged side chains in the e- and g-positions as well. Studies of side chain rotamer preferences in the hydrophobic core indicated that Leu and its hexafluorinated analogue prefer different conformations.

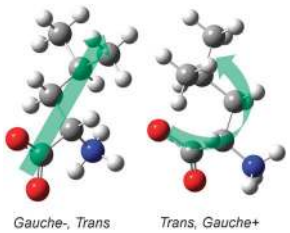


Fig. 6 Preferred side chain rotamers for Leu and 5^3,$5'^3$-F_6Leu. The *trans*, gauche + rotamer is disfavoured due to steric clashes and electronic repulsion for 5^3,$5'^3$-F_6Leu (reproduced with permission from Pendley *et al.*[45] Copyright 2011, John Wiley and Sons).

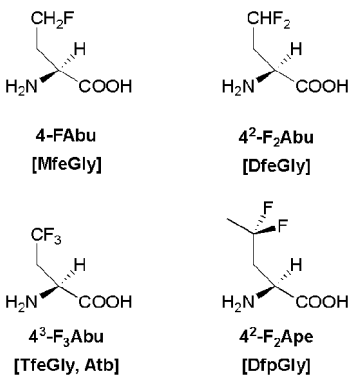


Fig. 7 Structures of α-L-aminobutyric (Abu) and α-L-aminopentanoic acid (Ape) analogues that have been used to replace single Leu residues in coiled-coil model peptides.

The increased steric bulk of the δ-CF_3 groups apparently evoke steric clashes between the fluorine atoms and the backbone. Additionally the C–F bond exhibits a negative partial charge as opposed to the C–H bond and these negative partial charges might come in close contact to the negatively charged backbone oxygen in the less favoured *trans*, *gauche* + rotamer (Fig. 6). These findings suggest that fluorinated analogues of aliphatic amino acids cannot be considered fully isomorphic with their hydrocarbon counterparts.

All of the studies reviewed above mainly addressed the effects of multiple or global replacements of hydrophobic amino acids by highly fluorinated analogues in coiled-coil models. In these models fluorinated amino acids for the most part interact with fluorinated amino acids in the hydrophobic core. Our own group focused on probing the interactions with native amino acids using two different models, model I an antiparallel homodimer,[46,47] and model II a parallel heterodimer[48,49] that contained single substitutions within the coiled-coil monomers (Fig. 8). Both of these models were designed to guarantee that the fluorinated side chain exclusively interacts with native hydrophobic amino acids. By incorporating analogues of α-L-aminobutyric (Abu) and α-L-aminopentanoic acid (Ape, Fig. 7) both models were used to probe the effects of a gradual increase in fluorine content on these coiled-coil interactions. Although these residues are not necessarily analogues of canonical L-amino acids, their application in protein engineering may allow for general conclusions.

These studies were initially based on the assumption that a CF_3 and an isopropyl group would be comparable regarding steric effects. 4^3-F_3Abu as a substitute for Leu, however, did not yield equally stable coiled-coils due to its lower hydrophobicity and different shape. Moreover, the effects of single amino acid substitutions were shown to be strongly position dependent. Most pronounced differences have been observed for 4^2-F_2Ape. While in the antiparallel model 4^2-F_2Ape strongly destabilizes the assembly when present at an a-position, it yields the most stable fluorinated parallel heterodimer. In fact, at position a of the parallel coiled-coil heterodimer stability largely correlates with the size of the side chain. However, such a clear correlation between size and stability was not observed for the variants modified at the d-position. By comparison with crystal structures and MD simulations it was concluded that the different side chain packing at the a- and d-positions must have a pronounced impact on their net effect on stability (Fig. 9). While at the a-position the strongly polarized β-methylene groups point away from the hydrophobic core, they point into it at position d thereby disturbing hydrophobic interactions.

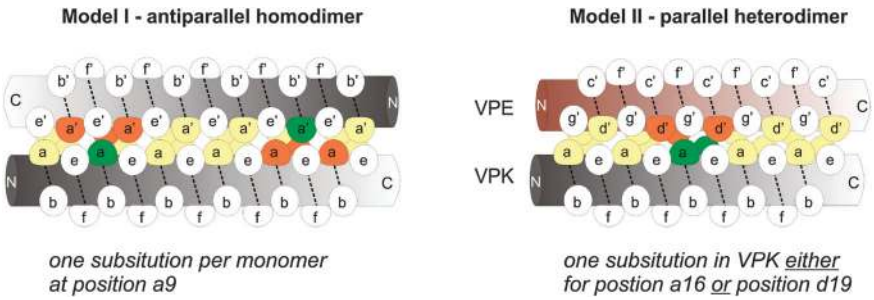


Fig. 8 Schematic representation of the coiled-coil models that have been used to evaluate single substitutions applying analogues of Abu.

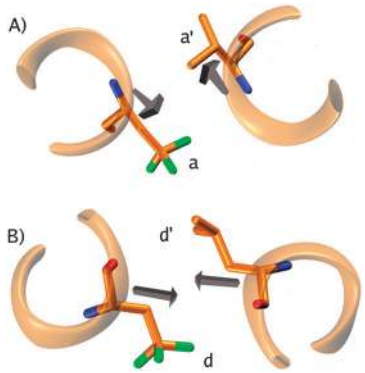


Fig. 9 Snapshots from MD simulations show the different packings of a fluorinated side chain (4^3-F_3Abu) with its directly interacting partner within the parallel coiled-coil model system at (A) position a and (B) position d (reproduced with permission from Salwiczek *et al.*[48] Copyright 2011, John Wiley and Sons).

Thus, the effect of partly fluorinated side chains depends on their packing/orientation within the core. Investigations of the association kinetics[50] have ultimately led to the hypothesis that a high degree of fluorination may affect coiled-coil association by the fluorous effect. Non-specific clustering of the bulky fluorinated head groups could be trapping the monomers in the unfolded state thus reducing the association rate. In general, the association and dissociation kinetics correlate with hydrophobicity and/or size with the exception of 4^3-F_3Abu. It may be interesting to include more heavily fluorinated amino acids such as $5^3,5'^3$-F_6Leu in these studies. Nevertheless, this hypothesis is hard to prove by experiment since such non-specific fluorous "interactions" would be difficult to detect due to their presumably low kinetic stability.

Other helical structures. Meng *et al.* have used $5^3,5'^3$-F_6Leu to modify the helical glucagon-like peptide 1 (GLP-1[7-36], H_2N-H**AEG**TFTSDVSSYLEGQAAKE**FI**AW**L**VKGR-$CONH_2$) first to modulate binding to its cognate human receptor (GLP-1-1R) and second to stabilize it towards proteolytic degradation by dipeptidyl peptidase IV (DPP IV).[51] The cleavage site for proteolysis is located between Ala8 and Glu9. Thus, Ala8, Glu9 and Gly10 were individually substituted. Furthermore, a double mutant containing the fluorinated residue at positions 8 and 9 was synthesized. All fluorinated variants were more stable towards cleavage by the protease. In fact, the Ala8-$5^3,5'^3$-F_6Leu as well as the Ala8Glu9-$5^3,5'^3$-F_6Leu double mutant completely resisted degradation. These findings are important in light of the fact that GLP-1 holds great promise as a treatment for type 2 diabetes, yet the native peptide has a half life of less than 2 minutes. In addition to these variants analogues that contain substitutions at sites that are crucial for receptor binding (Phe28, Ile29 and Leu32) were produced. All the fluorinated variants exhibited a reduced affinity towards the receptor. Most of fluorinated analogues were also active in stimulating cAMP production albeit at reduced efficiency because of their lower binding affinity. The finding that substitution of Glu9 by $5^3,5'^3$-F_6Leu had the least unfavourable effect on binding affinity may initially appear surprising since the CF_3 groups are not charged. However, there appears to be sufficient space in the binding pocket for such bulky residues. Another very interesting hypothesis is that the CF_3 groups may engage in multipolar C–F···C=O, C–F···H–X (X=O, N, S), C–F···H–C^{α} and C–F···side chain interactions with Arg, Gln and Asn.[52]

Employing *in vivo* procedures Wang *et al.* substituted diastereomeric 5^3-F_3Ile for five conserved isoleucines within the helical core of the cytokine murine interleukin-2 (mIL-2), a common anti-cancer drug.[53] The efficiency of their incorporation was tested beforehand using murine dihydrofolate reductase as a model protein. Unfortunately, none of the $3'^3$-F_3Ile diastereomers supported protein biosynthesis even under conditions in which isoleucyl-tRNA-synthetase was over-expressed. The fluorinated mIL-2 was analysed for activity by measuring the proliferative response of IL-2 dependent H2-T cells. The concentration at which the protein elicits 50% of the maximal proliferative response was about 30% higher for the fluorinated variant (EC_{50} = 3.87 ng mL^{-1} compared to EC_{50} = 2.70 ng mL^{-1} for wild type mIL-2). Based on the finding that the mutant was almost as active as the wild type protein, the authors argued that the fluorinated mIL-2 must fold into an authentic native-like structure. Direct structural investigations, however, have not been pursued.

A property that is often overlooked when interpreting the results of fluorinated amino acid substitutions is their actual propensity for a certain secondary structure. Indeed, it has been found that fluorinated amino acids such as analogues of Leu, Abu and Phe exhibit a reduced α-helix propensity compared to their hydrocarbon analogues.[54,55] This conclusion has been drawn from investigations of a monomeric α-helix Ac-YGG-KAAAAKA**X**AAKAAAAK-NH_2 that served as a host–guest model into which 4^3-F_3Abu, $5^2,5'^2$-F_4Leu, $5^3,5'^3$-F_6Leu and 2,3,4,5,6-F_5Phe were incorporated at the X-position. While the lower helix propensity may in part result from unfavourable solvent interactions at exposed positions of the monomeric helix, it is not clear as to what extent helix propensity affects protein stability at buried positions, for example, within the hydrophobic core of a coiled coil. Based on the finding that Abu has a higher helix propensity than 4^3-F_3Abu, yet destabilizes a coiled-coil dimer,[48] we have argued that helix propensity may play a minor role in such cases. Moreover, until all fluorinated amino acids have been investigated in this respect, general conclusions cannot be made.

2.3.2 β-Sheet structures. It has been argued that since some fluorinated amino acids exhibit low helix propensities, they may be more suitable for accommodation within β-sheet domains. Horng and Raleigh substituted diastereomeric 4^3-F_3Val for either Val3 or Val21 within the β-sheet core structure of the N-terminal domain of the ribosomal protein L9 (NTL9[1–56], Fig. 10).[56] Although the per-residue increase in stability of 0.8 kcal mol^{-1} at position 3 and 1.4 kcal mol^{-1} at position 21 refers to NTL9 variants that contain mixtures of both diastereomers, it was larger than for any helical protein. The kinetics of folding and unfolding revealed that fluorination increases the folding rate while decreasing the unfolding rate. Both of these observations may be explained by the increase in hydrophobicity upon fluorination that stabilizes the native state and destabilizes the unfolded state relative to the transition state.

Interestingly, the stabilization of the protein by single Val to 4^3-F_3Val substitutions was sufficient enough for including

MKVIFLKDVKGKGKKGEIKNVADGYANNFLFKQGLAIEATPANLKALEAQKQKEQR

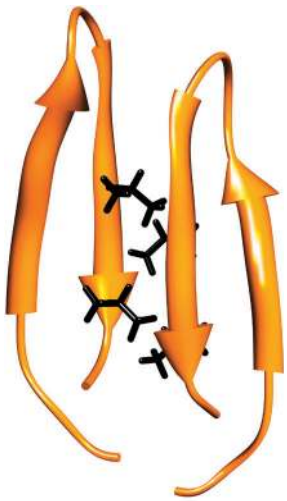
Fig. 10 Structure of NTL9(1–56) with highlighted residues Val3 and Val21 (PDB code 1DIV). The figure has been generated using UCSF Chimera.[57]

these analogues in an extensive fine structure analysis of the NTL9(1–56) folding transition state.[58] These more recent studies also included 5^3-F_3Leu as a substitute for Leu30 within the short α-helix. Here, the CF_3 group destabilized the protein, presumably due to steric clashes within this tightly packed region of the protein. Although the reported stability differences may allow for such investigations, a fine structure analysis (Φ-analysis) should employ stereochemically pure amino acids. To our understanding, however, both 4^3-F_3Val and 5^3-F_3Leu were incorporated as mixtures of two diastereomers with respect to C^β and C^γ, respectively. Therefore, interpretations on the level of almost atomic resolution should be made carefully. A comparison of the individual diastereomers may provide even deeper insight.

Since Horng *et al.* studied two buried positions of NTL9(1–56), it has remained unclear as to whether the observed stabilization was due to more favourable hydrophobic interactions, an increased propensity of the fluorinated amino acid for β-sheet conformations or a combination of both. Therefore, Chiu *et al.* used the GB1 domain as a model to study the intrinsic propensity of fluorinated amino acids (4^3-F_3Abu, 4^2,$4'^2$-F_4Leu, 5^3,$5'^3$-F_6Leu as well as F_5Phe) for β-sheet conformations by substituting the solvent exposed internal β-strand position 53 (Fig. 11).[59] All fluorinated amino acids were shown to increase the stability of GB1 compared to their non-fluorinated counterparts. This was attributed to increased hydrophobicity as well as steric effects. The authors argued that large side chains can favour sheet conformations due to limitations of the available backbone conformations. However, the stabilization was less pronounced than when observed by Horng and Raleigh upon introducing 4^3-F_3Val into NTL9.[56] At a solvent exposed position the more

MTYKLILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWTYDDATKTFTVTE

Fig. 11 Crystal structure of GB1 highlighting position 53 (PDB code 1PGA).[60] The figure has been generated using UCSF Chimera.[57]



RGGRLCYCRRRFCVCVGRX

Fig. 12 Solid-state NMR structure of the membrane-bound protegrin-1 dimer (PDB code 1ZY6)[62] with highlighted valines in positions 14 and 16 of each monomer. The figure has been generated using UCSF Chimera.[57]

hydrophobic fluorinated amino acids would unfavourably interact with the aqueous solvent thus partly attenuating the stabilizing effect. In conclusion, these studies support the notion that at least the fluorinated amino acids investigated in this study have a higher β-sheet propensity.

The relationship between the biophysical properties and biological activity of a membrane active β-hairpin, protegrin-1 (H_2N-RGGRLCYCRRRFCVCVGR-$CONH_2$), was examined by Gottler *et al.*[61] Two important Val residues at position 14 and 16 (Fig. 12) within the hydrophobic core of the dimer were replaced by Leu and 5^3,$5'^3$-F_6Leu. Both substitutions increased the hydrophobicity of the peptide as judged from their retention times, whereby the fluorinated peptide was the most hydrophobic as expected. Severe perturbations of the protegrin-1 structure were not seen, both in the absence and presence of lipid membranes. It has been reported that the incorporation of highly hydrophobic fluorinated amino acids may inappreciably alter haemolytic activity of helical AMPs (as is the case for magainin-2, see Section 2.3.1). All variants of protegrin-1 showed equal haemolytic activities. However, while the Leu variant showed enhanced antimicrobial activity against several bacteria, the fluorinated variants mostly appeared less potent. Thus, a simple correlation between antimicrobial activity and hydrophobicity could not be concluded. Most importantly, the Leu and 5^3,$5'^3$-F_6Leu substitutions changed the apparent stoichiometry of the peptide lipid interaction.

The authors suggested that instead of forming dimers both the Leu and the hexafluoroleucine variant formed tetramers upon binding to the membrane (Fig. 13). Their observed binding isotherms suggest that the fluorinated variant exhibits a lower effective positive charge which may explain its lowered activity (the initial binding to the negatively charged membranes is mediated by the high positive charge of protegrin-1).

2.3.3 Globular proteins. The preceding sections deal with substitutions within defined secondary structural elements of

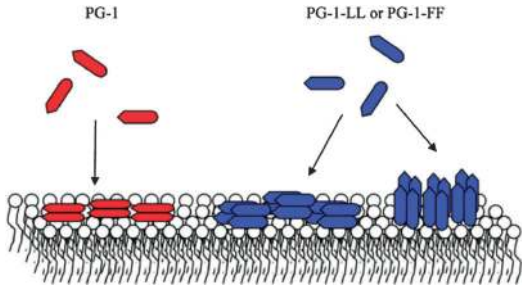


Fig. 13 Schematic binding pattern of native protegrin-1 and the variants containing Leu (PG-1-LL) and $5^3,5'^3$-F_6Leu (PG-1-FF, reproduced with permission from Gottler *et al.*[63] Copyright 2011, American Chemical Society).

peptides and proteins. In this section substitutions are addressed that were made in less well defined regions of proteins. Global substitutions of aliphatic residues that are dispersed throughout the whole structure of globular proteins are considered as well.

Alexeev *et al.* studied different site-specific (4*S*)-5-FLeu substitutions for Leu within the hydrophobic core of ubiquitin (Ub).[64] Selected pairs of Leu residues, Leu43/Leu67 and Leu50/Leu67, that are in close contact within the hydrophobic core were chosen to assess the utility of (4*S*)-5-FLeu as a ^{19}F-reporter. The overall structure of the fluorinated variants was comparable to the native protein as judged from their CD spectra.

The Leu50/67(4*S*)-5-FLeu mutant was chosen for further investigation by means of X-ray structural analyses as well as thermodynamic analyses. Although the thermal stability of Ub-Leu50/67(4*S*)-5-FLeu is slightly lower, its crystal structure (Fig. 14) and that of native Ub are practically superimposable. It is interesting that the position of the C-terminal loop was shifted. Since the C-terminus is known to be flexible and far away from the substitution, this might be a crystallization artefact. The authors shortly discussed probable causes for the loss in stability upon substitution. Since the melting point merely reflects the ratio of the enthalpy and entropy of unfolding and the latter has not been reported, however, it is not clear to what extent and why the protein has been destabilized by means of changes in free energy. Since monofluorination tends to decrease hydrophobicity, one may hypothesize that the presence of the supposedly more hydrophilic monofluorinated leucine within the hydrophobic core destabilizes the protein.

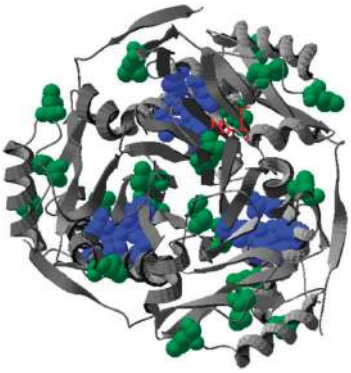
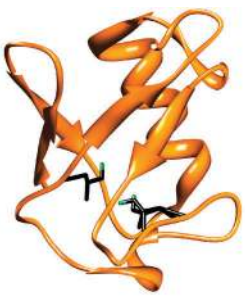


Fig. 15 Structure of CAT with highlighted leucines. Leucines that are in the hydrophobic core are marked in blue, the other leucines are coloured in green (reproduced with permission from Panchenko *et al.*[65] Copyright 2011, John Wiley and Sons).

Several studies have shown that the global substitution of fluorinated analogues for hydrophobic amino acids in globular proteins, as opposed to the same substitutions in the hydrophobic core of coiled coils, for example, tends to reduce stability. Panchenko *et al.* presented a study in which they substituted 17, 37 and 80% of all leucines of chloramphenicol acetyltransferase (CAT, Fig. 15) by diastereomeric 5^3-F_3Leu applying *in vivo* methods.[65] Although maximal fluorination appeared to enhance secondary structure formation, it was shown that the stability of the trimeric protein decreased along with the extent of substitution. Moreover, misaggregation of the protein was found at elevated temperatures. The authors picked up a speculative conclusion made by Marsh earlier.[66] Based on the concept that "fluorines prefer to exist near other fluorines", *i.e.* the fluorous effect, misfolding may be caused by fluorination of exposed residues that could form clusters in non-native folding states.

Similar results were found for the green fluorescent protein (GFP) which lost its fluorescence upon global substitution of diastereomeric 5^3-F_3Leu for Leu.[67] The reason was misfolding and aggregation of the fluorinated variant. Further investigations on both CAT and GFP, however, revealed that it is possible to evolve fluorinated variants that exhibit enhanced folding and function.[67,68] Protein expression in the presence of 5^3-F_3Leu combined with directed evolution[69] resulted in a fluorinated CAT variant with improved thermal stability.[68] A comparable study led to the evolution of a fluorinated GFP



MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQ**X**EDGRTLSDYNIQKEST**X**HLVLRLRGG

Fig. 14 Crystal structure of the Ub-Leu50/67(4*S*)-5-FLeu mutant (PDB code 1OGW).[64] The figure has been generated using UCSF Chimera.[57]

mutant that exhibited a 650-fold enhanced fluorescence.[67] In summary, stabilization of protein interactions and structure by global fluorination can be achieved, when the fluorinated aliphatic residues constitute distinct interaction cores of the protein. While this mostly applies to homotypic interactions, most protein ligand interactions that are of significance for drug development are heterotypic. Moreover, globular proteins usually do not tolerate global fluorination and require reorganization, *i.e.* evolution of their primary structure to compensate for the loss in stability and activity.

2.4 Methionine

Methionine can participate in hydrophobic interactions and hydrogen bonding or metal coordination. The presence of fluorine atoms in the S-methyl group can alter the size and shape of the methyl group. Also it reduces the electron density at the sulfur nucleus thereby increasing the overall hydrophobicity of the side chain.[70] Although other fluorinated methionines have been reported as well, relevant studies of proteins exclusively deal with S-fluoromethylated analogues (Fig. 16).

A more practical concern in protein analytics is that, due to the reduced nucleophilicity of the sulfur atom, peptide bonds involving TfM are not susceptible to cyanogen bromide mediated cleavage.[71] Whether the reactivity actually correlates with the extent of fluorination remains unclear, because MfM and DfM have not been investigated in this respect.

DfM and TfM were used by Houston Jr. *et al.* to modify the chemotactic tripeptide fMet-Leu-Phe.[72] They found that the fluorinated peptides show maximal neutrophil responses at a ten-fold lower concentration compared to the non-fluorinated parent peptide. The increased chemotactic activity was discussed to result from the increased size and hydrophobicity of fluorinated methionine that both contribute to fitting the side chain into the hydrophobic pocket of the receptor. Preliminary theoretical investigations have indicated that the presence of fluorine in methyl thioethers may also alter the preferred conformation. Explicit structural data were unfortunately not available.

DfM and TfM have also been incorporated into bacteriophage lysozyme (LaL) by means of protein expression using Met auxotrophic bacterial strains to probe their potential as ^{19}F-reporters.[73–75] Protein biosynthesis has been supported to different extents by these analogues. While the level of incorporation reaches up to 95% for DfM, the maximum level achieved for TfM was only 70%. In general, however, fluorination did not compromise the enzyme's activity. The ^{19}F-NMR investigations indicated that fluorination may induce very slight structural perturbations such as a reduction in rotational freedom of the S-methyl group. Due to the high sensitivity of the ^{19}F nucleus such subtle changes can be detected. Therefore, fluorinated Met has proven a useful spectroscopic probe for protein structure analysis.

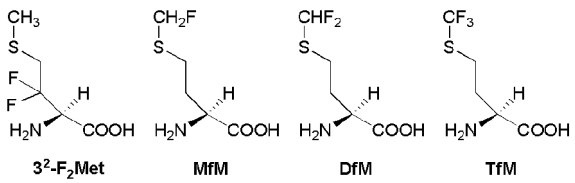


Fig. 16 Structures of fluorinated L-methionine analogues for which syntheses have been described.

Budisa *et al.* reported an analogue of enhanced green fluorescent protein (EGFP) that contained two Met to TfM substitutions (all other Met residues were mutated), one buried within the core (position 218) and one exposed to the solvent (position 1)[76] However, the levels of incorporation were rather poor but they found that the N-terminal TfM was not removed after protein biosynthesis. Posttranslational processing of the N-terminus has proven to be vital to cell growth in certain cases. The authors pointed out that formylated TfM may thus provide the basis for developing tumor therapies.

Generally, methionine residues share a rare abundance in naturally occurring proteins. Thus, it may be less surprising that substitutions by homologous fluorinated analogues have only marginal effects on the structure and stability. Some proteins contain conserved methionines whose structural and functional importance has yet to be clarified. For example, alkaline protease A (AprA) from *Pseudomonas aeruginose*, a member of metzincin superfamily of metalloendoproteases, includes a conserved Met-β-turn in proximity to the active site. Walasek and Honek studied the impact of substitution by DfM and found a marginal impact on structure and function.[77] In agreement with the findings by Budisa *et al.*[76] they observed that the N-terminal Met, which initiates biosynthesis, is less prone to posttranslational removal when substituted by DfM. Therefore, while DfM seems to affect binding to Met-amino-peptidase, it only slightly reduces the catalytic efficiency of the protease itself. Moreover, the thermal stability was practically identical to that of the wild type.

DfM and TfM were also substituted for a conserved methionine in the copper binding centre of azurin, an oxido-reductase, to investigate their impact on the reduction potential.[78] It was found that, although fluorination strongly reduces the thioether's nucleophilicity, both DfM and TfM can still sufficiently participate in metal coordination. While the spectroscopic properties (UV and EPR) of the protein have been largely unaffected by the mutations, the reduction potential exhibits significant alterations. The authors studied many copper binding proteins and concluded that the reduction potential does not correlate with the strength of coordination but rather with the hydrophobicity of the coordinating ligands.

One protein that contains a large number of Met residues is *KlenTaq* polymerase of *Thermus aquaticus* (14 positions including the N-terminus). Holzberger *et al.* have generated a globally fluorinated analogue in which up to 82% of all Met positions were substituted by TfM.[79] Despite a marked decrease in thermal stability, the error rate during replication did not alter.

3. Analogues of aromatic amino acids

3.1 Phenylalanine

3.1.1 General properties. Substituting hydrogen with fluorine on the aromatic ring of phenylalanine leads to a rearrangement of the electrostatic potential and increases the hydrophobicity of the aryl side chain. While single hydrogen

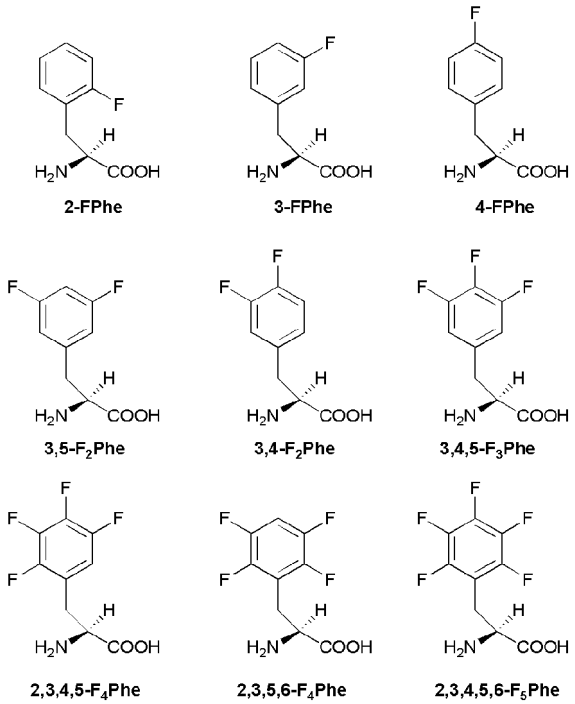


Fig. 17 Structures of Phe analogues with fluorine substitutions at the aromatic ring.

substitutions cause negligible perturbation,[80] the increased size of pentafluorophenylalanine may lead to destabilization of proteins owing to steric repulsion (molecular volumes of C_6H_6 and C_6F_5H are 106 $Å^3$ and 141 $Å^3$).[81] However, the perfluoro-substitution of phenylalanine results in a quadrupole moment of opposite polarity, which may enhance attractive interactions between fluorinated and non-fluorinated phenyl rings.

Fig. 17 depicts all phenylalanine variants that have been incorporated in line with the here reviewed studies. The syntheses of the C^{β} fluorinated analogues have been described,[82,83] yet they have never been incorporated into peptides or proteins.

3.1.2 General effects on protein stability

Incorporation of monofluorinated phenylalanine isomers. The three monofluorinated analogues of Phe (2F-, 3F-, and 4F-Phe) bear the same size and overall polarity but alter in side chain shape and dipole direction. Thus, they can be used to investigate distinct positional effects of fluorine substitution on aromatic side chains in biologically relevant globular proteins.

Global substitution of phenylalanine with 2-, 3-, or 4-FPhe in human annexin V and *P. aeruginosa* azurin has been accomplished by applying selective pressure incorporation (SPI).[84] Bioincorporation of the fluorinated phenylalanine analogues was analysed by UV spectroscopy as well as mass spectrometry and structural integrity was confirmed by CD spectroscopy. Interestingly, the proteins containing fluorinated phenylalanines exhibit two characteristic shoulders (fingers) in the UV spectra in the range 260–270 nm, and thus provide a way to use these building blocks as spectroscopic protein labels.

The three monofluorinated phenylalanine isomers were also globally incorporated into *Pvu*II restriction endonuclease by means of SPI.[85] Each *Pvu*II subunit features four Phe residues outside the active site. Incorporation levels of 2F-, 3F- and 4F-Phe were 7%, 17%, and 8%, respectively. Heterogenous mixtures of fluorinated mutants were used for subsequent investigations. Even though the 3F-Phe mutant exhibited a similar stability towards guanidine hydrochloride denaturation as its wild type, it possesses a two-fold higher average specific activity. The 2F-and 4F-Phe mutants, on the other hand, exhibited a decrease in specific activity and conformational stability. These results highlight the impact of subtle conformational changes on enzyme activity, since none of the investigated positions are part of the DNA recognition or catalytic site.

Moreover, the three monofluorinated phenylalanines have been globally substituted for the 10 Phe residues of histone acetyltransferase (HAT) tGN5 applying SPI.[86] Based on MALDI MS of the peptides, the extent of phenylalanine replacements was calculated to be above 80% for all three analogues. A loss in the secondary structure as well as a decrease in thermal stability was monitored for all three mutants by CD experiments. Furthermore, contrary to the widely held belief that incorporation of fluorinated amino acids enhances protease resistance, digestion with chymotrypsin (which cleaves after Phe) revealed that the presence of any of the monofluorinated phenylalanine isomers leads to reduced enzymatic resistance when compared to the wild type.

Recently, high level substitution of all 11 *Candida Antarctica* lipase B (CalB) phenylalanines with 4F-Phe was achieved through SPI in yeast, as confirmed by mass analysis.[87] However, the incorporation was stochastic and yielded a mixture of fluorinated proteins. CD spectroscopic analysis revealed that the secondary structure profile remained intact after substitution by 4F-Phe, yet lowered intensities signified a destabilization relative to the canonical lipase. The resistance towards proteinase K, which cleaves after aliphatic and aromatic amino acids, remained unchanged. Furthermore, the catalytic activity of the fluorinated enzyme mutant was lower than that of the parent protein, but phenylalanine fluorination appeared to prolong the shelf life of CalB as it showed enhanced activity after refrigerated storage over several months when compared to its wild type.

Altogether, investigations on the positional effects of monofluorinated phenylalanines reveal that even subtle positional changes of a single fluorine can influence the structure and function of enzymes dramatically.

Incorporation of pentafluorophenylalanine (2,3,4,5,6-F_5Phe; F_5Phe). Investigations on the helix stability within a Ala-Lys model peptide identified 2,3,4,5,6-F_5Phe (F_5Phe) as a helix breaker in the first instance, but an *i*, *i* + 4 Phe-F_5Phe arrangement resulted in an increasing overall helicity by a factor 2 relative to a *i*, *i* + 5 spaced control peptide.[88] Studies by Cheng *et al.* on the helicity of monomeric alanine-based model peptides have been able to corroborate the observed decrease in helicity upon replacing Phe with F_5Phe.[54] Substitution at a solvent exposed position of a monomeric β-sheets protein domain (G B1) resulted in an increased stability.[59]

Senguen and co-workers used F_5Phe to evaluate its aromatic, hydrophobic and steric effects on the self-assembly of the amyloid-β-fragment Aβ(16–22).[89] This short peptide segment (Ac-KLVFFAE-NH_2) contains two central phenylalanine residues, and is known to pack in antiparallel β-sheets within the

amyloid fibril. F_5-Phe was substituted for the two phenylalanines individually as well as simultaneously. Kinetic and thermodynamic investigations showed that both singly substituted F_5Phe variants self-assembled at dramatically enhanced rates relative to the wild type, and the double mutant exhibited additive thermodynamic effects. TEM images revealed ordered fibrils that are several μm in length, and X-ray powder diffraction of fibril pellets confirmed a characteristic amyloid diffraction pattern. Elevated fibril formation in fluorinated Aβ mutants was not only attributed to an increased hydrophobicity, but also to additional fluorine–fluorine interaction, since highly fluorinated amino acids show strong self-aggregating properties.[66] Additionally, a combination of aromatic effects (π-stacking of aromatic side chains) and a potentially higher β-sheet propensity of F_5Phe, as observed by Cheng *et al.*,[59] were proposed to contribute to enhanced assembly rates.

It was recently reported that the presence of F_5Phe in a tripeptide epoxyketone proteasome inhibitor has a pronounced effect on inhibitory potency and selectivity.[90] Proteasomes contain three proteolytically active sites (β1, β2, and β3) that feature different substrate specificities. The caspase-like β1 subunit cleaves after acidic residues, the trypsin-like β2 subunit after basic residues, and the chymotrypsin-like β5 subunit cleaves after bulky, hydrophobic residues.[91] It was found that the effect of incorporation of perfluorophenylalanine in the investigated proteasome inhibitor depends on the site of substitution.[90] Substituting the P2 and P3 sites dramatically decreases the potency of the inhibitor, while a substitution at the P2 site hardly affects the potency. Moreover, substituting in P2 generates a highly specific inhibitor of the proteasome's β5 subunit, that was also successfully converted to a β5 selective fluorescent probe (Fig. 18).

3.1.3 Modulation of aromatic interactions

Quadrupole interactions. Interactions between aromatic side chains involve hydrophobic, van der Waals, and electrostatic forces.[92,93] Hydrophobic and van der Waals forces lead to strong interactions with other hydrophobic or aromatic side chains, while the electrostatic component is assumed to contribute geometric preference to the interaction. The quadrupole moment of aromatic rings arises from the positively charged σ-framework between two regions of π-electron density on the faces of the ring.[94] The electrostatic potential of aromatic compounds is usually negative inside the aromatic ring, while the hydrogen atoms on the outside are partially positively charged. As a result, a face-to-face stacking

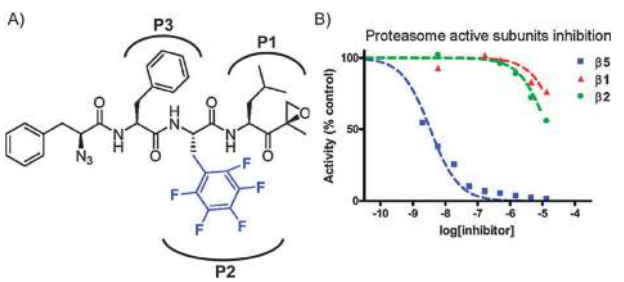


Fig. 18 (A) Structure of peptide epoxyketone proteasome inhibitor. Indicated are enzyme pockets (P1, P2, P3); (B) subunit activity post inhibition with proteasome inhibitor (adapted with permission from Guerink *et al.*[90] Copyright 2011, American Chemical Society).

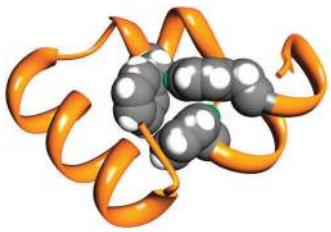


Fig. 19 Structure of villin head piece N68H mutant (PDB code 1YRF)[97] highlighting the hydrogens directly interacting with the phenylalanine electron density in light green.

arrangement of two aromatic rings is repulsive, whereas edge-to-face interactions are favoured.

One example involving quadrupole interactions of amino acid side chains is given by the villin headpiece subdomain (VHP, Fig. 19), a 35-residue three helix bundle structure comprising a hydrophobic core, which features a cluster of three edge-to-face oriented phenylalanines (residues 6, 10 and 17). In order to be able to compare conformational stabilities for all the possible Phe/F_5Phe combinations within the hydrophobic core, backbone thioester exchange (BTE) was applied by Gellman and coworkers.[81] To this end, a central amide linkage was replaced by a thioester (α-thio-acid residue), and the thioester–thiol exchange (equilibration between full length peptide and peptide fragments) was subsequently monitored. The equilibrium constant (K_{BTE}) provided insight into the free energy of tertiary structure formation ($\Delta G_{fold/BTE}$), which was shown to linearly correlate with that obtained from guanidinium hydrochloride denaturation studies ($\Delta G_{fold/Gdn}$). Only one of seven investigated Phe VHP mutants (Phe10 → F_5Phe) proved to be more stably folded than the native peptide, and its NMR structure revealed very little variation in backbone conformation or side chain packing relative to the wild type in follow up experiments.[95] It is conceivable that the gain in hydrophobicity upon pentafluorophenylalanine introduction is not sufficient enough to compensate the loss of stabilizing edge-to-face stacking interactions, which results in destabilization of the other fluorinated mutants. Taking these considerations into account, two tetrafluorinated phenylalanine isomers (2,3,4,5-F_4Phe and 2,3,5,6-F_4Phe) have been used to substitute Phe residues of VHP in subsequent studies by Zheng *et al.* to obtain an enhanced hydrophobicity as well as to retain the possibility to undergo aromatic interactions with the remaining hydrogen atom.[96]

The calculated partial charges of the remaining hydrogen atoms show a substantial increase (Fig. 20). CD as well as NMR spectroscopy indicate that all investigated mutants adopt stable and native-like structures. Thermodynamic analysis revealed that 2,3,4,5-F_4Phe and 2,3,5,6-F_4Phe can contribute more favourably to VHP stability than 2,3,4,5,6-F_5Phe (12 °C in T_M and −1.1 kcal mol^{-1} in ΔG for Phe10 → 2,3,4,5-F_4Phe mutant), which is attributed to retained edge-to-face quadrupole interactions by the authors.

Aromatic and perfluoroaromatic molecules have a tendency to associate with each other, due to an electrostatic attraction which is induced by their reversed quadrupoles.[92,98,99] The quadrupole moment of perfluoroaromatic molecules is of opposite polarity, owing to the electron withdrawing effect

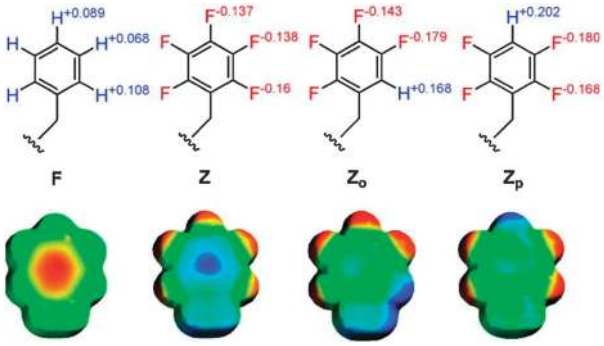


Fig. 20 Structures and space filling models depicting partial charges and electrostatic potentials of penta- and tetrafluorinated phenylalanine analogues. Negative and positive electrostatic potentials are indicated in red and blue, respectively (adapted with permission from Zheng *et al.*[96] Copyright 2011, American Chemical Society).

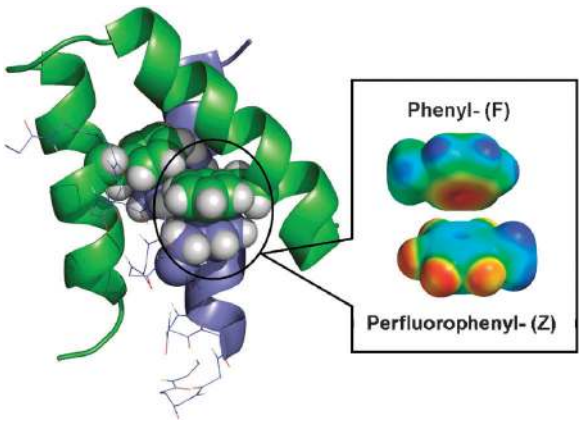


Fig. 21 Space-filling model highlighting electrostatic potentials of face-to-face stacked phenylalanine and perfluorophenylalanine side chains. Blue indicates a positive and red a negative electrostatic potential (reproduced with permission from Zheng and Gao[100] Copyright 2011, John Wiley and Sons).

of the fluorine atoms. This results in a net electrostatic attraction between parallel π-faces of aromats and perfluoroaromatic compounds, and thus in a fairly strong non-covalent interaction (Fig. 21).[100] The stabilization energy of benzene–hexafluorobenzene stacking interaction has been estimated to be 3.7 kcal mol^{-1} in gas phase theoretical studies.[99]

Experiments by Butterfield *et al.* revealed that Phe–F_5Phe interaction does not exhibit an enhanced stabilisation over the Phe–Phe interaction in an α-helical model peptide. Moreover, molecular modelling indicated an inability to access a fully stacked orientation of phenylalanine side chains in an α-helix.[88] However, Zheng and Gao recently demonstrated that an accurate positioning of aromatic and perfluoroaromatic rings can enable selective protein–protein interactions (Fig. 21).[100,101] The investigated model peptide α_2D folds into a homodimer that comprises a hydrophobic core with one phenylalanine residue from one monomer stacking face-to-face against a phenylalanine side chain of the other monomer. A double mutant in which both phenylalanines are replaced by the perfluorinated analogue also forms a dimer. However, it exhibits a tremendous increase in thermal stability (T_M = 80 °C as opposed to 30 °C of parental peptide), which is assumed to result from the increased hydrophobicity of the fluorine atoms in comparison to hydrogen atoms. When the two dimers are mixed they undergo a complete transition from homodimers to heterodimers (Fig. 22),[100] as confirmed by ^{1}H- and ^{19}F NMR studies as well as FRET experiments. The specificity of the heterodimer is expected to be an effect of the quadrupole interaction of stacked aromatic and perfluoroaromatic rings, whose electrostatic attraction was calculated to contribute as much as −1.0 kcal mol^{-1} to protein structural stability.

2,3,4,5,6-F_5Phe has also been used to enhance Phe/Phe interactions in thrombogenic collagen model peptides (CMPs) to facilitate their self-assembly into triple helix-based fibrils.[102] Collagens constitute a ubiquitous protein family in animals and humans that can help prevent bleeding and promote tissue repair in injured vessel walls by binding to the collagen binding receptor on platelets, which in turn induces platelet aggregation and thus blood clotting.[103] By attaching F_5Phe and Phe to the N and C termini of a $(GlyProHyp)_{10}$ collagen model peptide, respectively, Cejas *et al.* were able to demonstrate that aromatic stacking interactions can be used to encourage propagation by end-to-end stacking into fibrils similar to that of collagen.[102] XED force-field calculations on the energetics for head-to-tail stacking of triple-helical homodimers revealed the total binding energy to be −83.5 kcal mol^{-1}, which is significantly higher than that of the nonfluorinated phenylalanine analogue, with a calculated binding energy of −70.4 kcal mol^{-1}. Structural analysis of the fluorinated CMP was carried out by CD spectroscopy, TEM as well as light- and atomic force microscopy. Moreover, its ability to mimic collagen's biological function was assessed by a platelet aggregation assay. The peptide had a stable triple-helical character, formed micrometre-length fibrillar material that resembles collagen fibrils, and was potent in inducing platelet aggregation.

In summary, these results underline the necessity of suitable static geometries that provide for the correct stacking configuration of aromatic side chains, and furthermore show that the quadrupole interaction of the phenyl-pentafluorophenyl pair is strong enough to direct protein–protein interactions and fibril formation of collagen peptide mimics.

Cation–π interactions. A cation–π interaction describes the electrostatic attraction between a cation and the negative electrostatic potential associated with an aromatic π-face.[104] This non-covalent binding force is energetically comparable to a hydrogen bond. In proteins it occurs between aromatic and positively charged amino acid side chains (*e.g.* Lys or Arg).[105] Fluorination of aromatic amino acids reduces the electron density of the π-system, thereby its negative character, and thus potential cation–π interactions are weakened. Therefore, these amino acid analogues can be used to probe the involvement of cation–π interactions in folding and recognition processes of proteins and their ligands.[104,106]

In order to elucidate whether one or all three Phe residues of the glycine receptor binding site are involved in a cation–π interaction with glycine's amino group, 4-FPhe, 3,5-F_2Phe, and 3,4,5-F_3Phe have been site-specifically incorporated into this protein *via* nonsense suppression by Pless *et al.*[107] Only for one of the three Phe residues (Phe159) a strong linear correlation

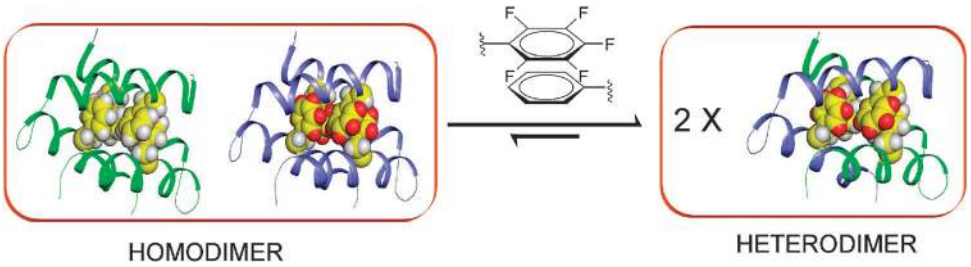


Fig. 22 Quadrupole stacking between aromatic and perfluoroaromatic rings leading to specific peptide–peptide interactions (reproduced with permission from Zheng and Gao[100] Copyright 2011, John Wiley and Sons).

between the decreased cation–π binding ability upon fluorination and receptor function could be observed. This indicates that a single cation–π interaction occurs during the binding process of the agonist to its receptor.

Recently, the same group undertook a similar study with glycine receptor agonists of lower efficacy, namely β-alanine and taurine.[108] Again, only Phe159 mutants showed a systematic decrease in agonist affinity upon an increasing degree of Phe fluorination, reinforcing the observation made before that a single cation–π interaction between glycine and its receptor has a strong impact on binding.

Tetrodotoxin (TTX) blocks sodium channels only when a Phe or Tyr residue is present at a specific site of this ion conducting pore.[109] To investigate the energetic contribution of a potential cation–π interaction to the binding process of the toxin, all three above mentioned fluorinated phenylalanine analogues have been incorporated at this key position (residue 401) in the rat skeletal muscle sodium channel $Na_v1.4$ *via* nonsense suppression.[110] With an increasing fluorine content of the aromatic ring a decreased affinity for TTX block was shown by the elevated TTX concentrations that are required to reduce the current in each mutant. Moreover, kinetic analyses revealed that upon successive fluorination the association rate constant (k_{on}) is decreased while the dissociation rate constant (k_{off}) is increased when Phe is substituted by 3,4,5F_3-Phe which is consistent with the progressive loss in TTX affinity.

Using a similar strategy the G protein-coupled receptor Ste2p was investigated to see whether a cation–π interaction is formed between the α-factor and Ste2p.[111] The binding of the α-factor to Ste2p leads to an activation of a signal transduction pathway that results in an arrest of cell growth. The C-terminal Tyr residue (Tyr13) of the tridecapeptide α-factor penetrates into the hydrophobic region of Ste2p and facilitates the binding to the receptor. Tyr13 of the α-factor was substituted with 3-FPhe, 4-FPhe, 3,4-F_2Phe, and 2,3,4,5,6-F_5Phe by SPPS in order to compare the agonist activities as well as equilibrium binding constants (K_d) of these analogues with the wild type.[111] All fluorinated derivatives retained some agonist activity, while wild type α-factor efficiently competed with their binding to the receptor. These findings indicated that these analogues interact with Ste2p. Furthermore, upon progressive fluorination of Phe a linear increase in binding affinity was observed. In fact, the K_d values correlate with the predicted energies of a cation–π interaction. Therefore, the authors concluded that these results are evidence for a cation–π interaction between α-factor's Tyr residue and an Arg in the hydrophobic region of Ste2p.

3.1.4 Modifying antigenic peptides. T cell receptor (TCR) recognition of antigenic peptides bound and presented by major histocompatibility complex (MHC) molecules is the basis for the cellular immune response. The binding process bears relatively low affinity and fast kinetics.

In order to investigate peptide–MHC interactions 2,3-F_2Phe was incorporated into a receptor binding position of a native nonapeptide from the Wilms tumor protein (WT1),[112] an antigen that is selectively overexpressed in leukemias and cancer cells and has been shown to generate a T cell response.[113] The 2,3-F_2Phe WT1 analogue was able to induce T cell response, but an enhancement in immunogenicity due to the increased hydrophobicity upon fluorination could not be observed.[112]

Applying a structure guided design approach the antigenic human T-cell lymphotropic virus Tax peptide (referred to as Tax) was modified with several fluorophenylalanine analogues, in order to selectively modulate TCR binding.[114] When presented on a MHC molecule, Tax is recognized by the A6 as well as the B7 T cell receptor.[115] In the antigen–TCR complex both receptors form a pocket that accommodates a tyrosine residue of the Tax peptide. In TCR A6 an Arg side chain acts as a hydrogen bond donor to tyrosines hydroxyl group, and in TCR B7 an Asp side chain introduces negative charge density, and thus acts as a hydrogen bond acceptor. It was found that substituting the central Tyr residue (position 5) in Tax with either 3-FPhe or 4-FPhe results in a 2-fold and 3-fold increased affinity towards TCR A6, respectively, and additive effects could be observed for the doubly substituted 3,4-F_2Phe variant (five-fold enhancement). While no binding with either receptor was monitored with the F_5Phe analogue, a substantial decrease in binding affinity with TCR B7 (up to 14-fold) was seen for all the investigated fluorinated peptides. SPR experiments undertaken with the 4-FPhe and 3,4-F_2Phe Tax variants showed a decreased dissociation rate for TCR A6 binding with both peptides, and ITC determined binding thermodynamics indicated that the affinity gain is entropically driven and enthalpically opposed. The X-ray crystal structures of these two analogues with TCR A6 are nearly superimposable with that of the native peptide, varying only in the region accommodating the substituted amino acid side chain. In both structures the Arg residue of TCR A6 has moved away from the fluorine atoms probably to optimize the distance for a favourable electrostatic interaction. However, a purely electrostatic mechanism can be expected to be enthalpically favoured. Due to the observed conformational changes and binding thermodynamics, the authors postulated a 'polar

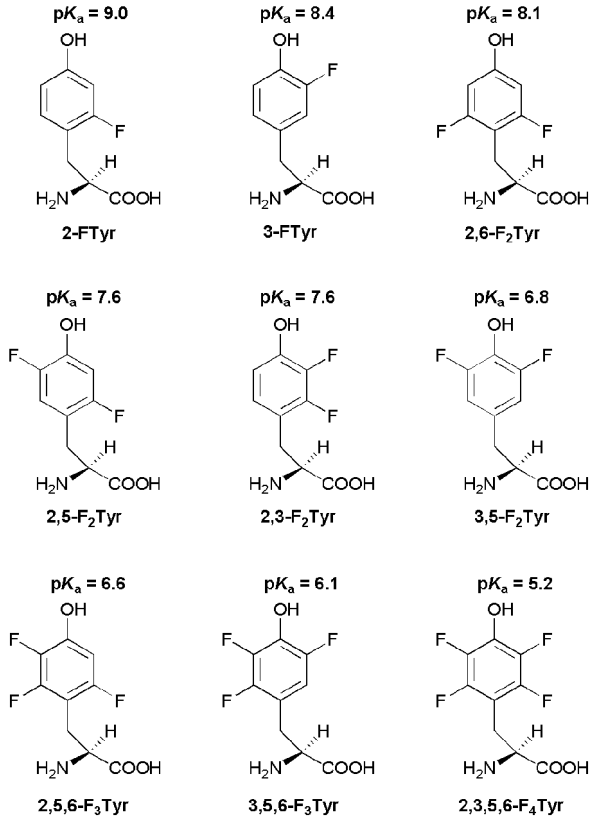


Fig. 23 Structures of L-tyrosine analogues with fluorine substitutions at the aromatic ring.

hydrophobicity' binding mechanism in which both electrostatic as well as hydrophobic interactions accounted for the enhancement in affinity.

3.2 Tyrosine

3.2.1 General properties. In theory, 12 fluorinated analogues of tyrosine are possible consisting of two diastereomers by single and one doubly substituted C^{β}-analogue as well as nine by substitutions of the aryl hydrogens. However, the literature exclusively deals with Tyr analogues that have been substituted at the aryl side chain (Fig. 23).

The strong electron withdrawing effect of fluorine leaves the aromatic ring more electron deficient. However, as opposed to fluorinated Phe, fluorinated Tyr analogues are more hydrophilic. This effect is due to the polarization of the hydroxyl group that becomes more acidic and a better hydrogen bond donor upon fluorination. The overlap of fluorine's lone pair electrons with the π-orbitals of the aromatic ring produces a moderate electron donating resonance effect (+M) which then redistributes the electrons and places more negative charge at the adjacent positions. Theoretically, the inductive and resonance effects may counteract each other. However, the inductive effect prevails and the lowering of the pK_a value of the hydroxyl group has been verified experimentally (Fig. 23).[116,117] Another important aspect of Tyr fluorination is its effect on the reduction potential that ranges from 705 to 968 mV, depending on the extent and position of fluorination, while native Tyr has a peak potential of 642 mV.[118] By manipulating the electronic environment of protein active sites by incorporation of fluorinated analogues, tyrosine's role in acid/base catalysis or in enzymes that involve intermediate tyrosyl radicals during catalysis can be investigated. It may be worth mentioning that when employing fluorinated Tyr analogues for biological applications, one has to consider that 3-FTyr and 2-FTyr were seen to antagonize tyrosine utilization competitively in several biological systems.[119,120] Furthermore, 3-FTyr is enzymatically converted in the cellular milieu to fluorinated intermediates of the tricarboxylic acid cycle (fluoroacetate and fluorocitrate) *via* the tyrosine pathway.[121,122] As a consequence, its administration in test animals has had lethal effects.[123] It is conceivable, though, that peptides or proteins substituted with these amino acids may serve as prodrugs for cancer treatment that induce intracellular fluorocitrate synthesis when converted into the active drug (free amino acid) at the target site.[124]

3.2.2 Fluorotyrosines as mechanistic probes of biological processes. Since fluorine substitutions on the aromatic ring have a profound impact on the phenolic pK_a, fluorinated tyrosines have been applied as probes to investigate the role of Tyr in different enzymatic mechanisms and other biological processes. As such, site-specific incorporation of 2-FTyr, 3-FTyr and 2,3,5,6-F_4Tyr at 3 different Tyr sites (positions 93, 190 and 193) in the agonist binding site of a nicotinic acetylcholine receptor *via* stop codon suppression helped examine the role of tyrosine's hydroxyl group.[125] Interestingly, 2-FTyr and 3-FTyr at position 93 produced slightly more potent analogues than the wild type.

To determine the degree to which side-chain hydrogen bonding stabilizes the folded state of staphylococcal nuclease (SNase)[126,127] and to investigate the catalytic properties of glutathione transferase (GST),[128] the same three analogues were site-specifically substituted for the central Tyr residues of these enzymes by *in vitro* suppression and stop codon mutation. Denaturation studies of the SNase mutants revealed a correlation between stability (K_{app}) and the pK_a of the phenolic OH-group thereby providing evidence that intramolecular hydrogen bonds with Tyr side chains stabilize the folded state of the protein relative to the unfolded state.[126] The results for GST suggested that the incorporation of the more acidic fluorinated tyrosines alter the protonation state of the enzyme's ground state. While it has been discussed that the thiolate of GST is normally stabilized *via* hydrogen bonding (Tyr–OH$\cdots^{-}$SG), the findings indicate that fluorination may alter the protonation state so that a tyrosinate anion interacts with a protonated glutathione (FTyr–$O^{-}\cdots$HSG). Theoretical investigations led to similar conclusions.[128] This change in protonation state may explain the substantial loss in catalytic efficiency upon fluorination. High resolution crystallographic studies furthermore revealed that global Tyr substitution with 3-FTyr does not disturb the overall folding pattern of rat glutathione transferase M1-1. Some local conformations were altered upon fluorination, however, and new hydrogen bonds potentially involving the fluorine atoms were observed.[129–131] Therefore, this activity may be a result of both the change in protonation state and altered local conformations.

By altering the reduction potential through incorporation of 2,3-F_2Tyr, 3,5-F_2Tyr, 2,3,5-F_3Tyr or 2,3,5,6-F_4Tyr into *E.coli*

ribonucleotide reductase, it was shown that Tyr acts as a redox active amino acid in this enzyme.[118,132] 2,3-F_2Tyr has furthermore been used as a probe for long range proton coupled electron transfer in the same enzyme.[133]

2-FTyr, 3-FTyr and 2,3-F_2Tyr have also been site-specifically substituted for a Tyr in the active site of Δ^5-3-ketosteriod isomerase to investigate its involvement in the catalytic mechanism.[134] While the other analogues effectuated a loss in catalytic efficiency, the substitution of Tyr by 2-FTyr produced a more efficient enzyme. 3-FTyr was furthermore site-specifically substituted for Tyr in the active site of aspartate aminotransferase to elucidate the catalytic mechanism.[135] In this enzyme Tyr stabilizes the binding of the PLP cofactor by a hydrogen bond. Despite the increase in acidity of the Tyr–OH by fluorination this interaction was weakened by this substitution. Thus, other factors such as sterics and conformation may come into play as well. Crystal structures of the enzyme mutant were not available.

The reassignment of Tyr codons to 3-FTyr during fermentation with Tyr auxotrophic *E.coli* under selective pressure led to fully fluorinated annexins and azurins.[124] Fluorination did not disturb the overall structural properties as confirmed by CD spectroscopy. Nevertheless, the melting temperatures revealed a slight loss in thermal stability for the predominantly α-helical annexin while an increase was observed in the case of azurin, which possesses a large sized β-sheet structure. Since no crystal structures were available for both these fluorinated proteins, it was not possible to elucidate the specific contribution of single Tyr side chains.

More recently, 3-FTyr was substituted for an active site Tyr of HIV-1 reverse transcriptase to elucidate its involvement in substrate interaction.[136] Altered chemical properties upon global incorporation of 3-FTyr in organophosphate hydrolase have been reported.[137] The fluorinated mutants of this enzyme showed an extended pH-optimum, which shifted to acidic pH, as well as an enhanced thermal stability in the alkaline pH region.

Despite the observation that the crystal structure of human manganese superoxide dismutase (MnSOD), in which Tyr positions were globally substituted with 3-FTyr, is closely superimposable with that of the non-fluorinated wild type enzyme,[138] the catalytic activity of Fluoro-MnSOD was significantly reduced.[139] The authors proposed that the unfavourable alteration of the kinetic properties were not directly caused by the 3-FTyr residues in close proximity to the active site metal, but rather by 3-FTyr residues more distant (>7 Å) from it.[139]

Site-specific substitution of Tyr with 3-FTyr by misacylated suppressor tRNAs and a read through of nonsense codons at the active sites of vaccinia virus DNA topoisomerase I,[140] and human topoisomerse I,[141] showed a reduced activity for the fluoroanalogues in comparison to the wild type enzymes in both cases. The observed pH dependency of enzyme activity is consistent with the general acid/base catalysis mechanism proposed for the enzyme.[142]

The investigation of the driving force for electron and proton transfer in Photosystem II (PSII), in which Tyr is involved in the formation of a radical pair,[143] was enabled by the global substitution of Tyr with 3-FTyr.[144] In this study, the activation energy of 3-FTyr oxidation was decreased by 110 meV, which correlates with the finding of a 90 meV stabilization energy of the phenolate form of 3-FTyr when compared to Tyr.[118]

3.2.3 Fluorotyrosines as substrates for kinases and phosphatases. All of the nine fluorinated Tyr analogues have been incorporated into model peptides to analyze the transition state and substrate selectivity of tyrosine kinase catalyzed phosphoryl transfer reactions to and from these fluorotyrosine model peptide substrates.[116,117] The findings support the proposal of a dissociative reaction mechanism with the neutral phenol acting as the preferred substrate nucleophile and verify that local residues surrounding a phosphorylated tyrosine can influence recognition. A few years later, expressed protein ligation enabled the examination of a semisynthetic protein containing a ring-fluorinated Tyr analogue as a kinase substrate.[145] Interestingly, in agreement with the previous studies on synthetic peptides,[116,117] the substrate was more efficiently phosphorylated. Thus, the pK_a drop in fluorinated Tyr analogues compared to Tyr induces very similar rate effects in peptide and protein substrate. However, the latter studies revealed that tyrosine rotamer preferences may be especially important within the context of proteins *versus* peptides, and thus highlight the value of analyzing protein substrate selectivity with actual protein substrates, while investigations on model peptides may produce misleading results. Nonetheless, studies on the catalytic mechanism of the insulin receptor kinase with a peptide substrate containing 2,3,5,6-F_4Tyr by steady state kinetics and X-ray crystallography of the enzyme–substrate complex support a reaction mechanism favouring a dissociative transition state for this enzyme.[146]

Recently, the synthesis of fully protected 3,5-F_2Tyr (Fmoc-3,5-F_2Tyr(*t*Bu)) has been reported and this building block was successfully incorporated into individual peptides and combinatorial peptide libraries by solid phase peptide synthesis.[147] The resulting peptides were susceptible towards tyrosine phosphatases (PTPs) and resistant to tyrosinase action, and are therefore optimal phosphatase substrates.

3.2.4 Incorporation into fluorescent proteins. Due to their spontaneous and cofactor-less fluorescence, autofluorescent proteins (AFPs) have emerged as highly valuable tools in the life sciences as genetically encoded reporters.[148,149] However, their application is limited by photobleaching of the chromophore or light induced damage of the surrounding biological medium.[150]

In order to circumvent these drawbacks in EGFP and EYFP, all Tyr residues in these proteins were globally replaced by 2-FTyr and 3-FTyr using Tyr auxotrophic *E.coli* strains to vary the electronic transition energies of the chromophore and to analyze differences in the optical properties of the mutants.[151–153] While the crystal structure of the 3-FTyr EGFP mutant suggests that the fluorine atom may act as a weak hydrogen bond acceptor, it also revealed that 3-FTyr populates two different rotamers.[151] Moreover, the fluorescence emission of this mutant was red shifted and its intensity was by 10% less than that of native EGFP. All the 2-FTyr variants, however, adopted only one confirmation in the solid

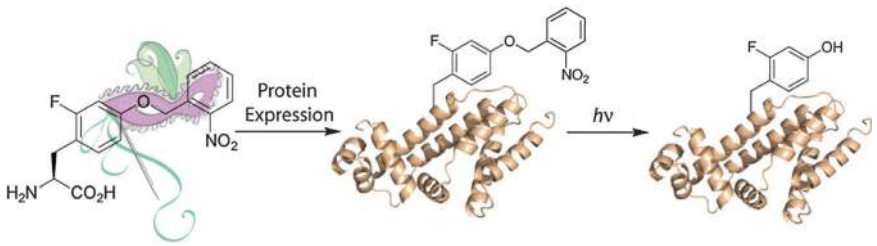


Fig. 24 Site directed incorporation of fluorotyrosines by photochemical disguise (reproduced with permission from Wilkins *et al.*[154] Copyright 2011, American Chemical Society).

state and the fluorescence emission maxima of the 2-FTyr variants are blue shifted.[152] A slight enhancement in fluorescence intensity for the EYFP variants was only observed at neutral pH and the photostabilty was not enhanced.[153] The authors assumed that Tyr fluorination led to a reduced shielding of the chromophore and thus to the formation of reactive oxygen species that may have been responsible for photobleaching.[150]

Photochemical disguise enabled the site specific incorporation of 3-FTyr and 2,6-F_2Tyr at the Tyr site of GFP that is central to the fluorophore (Fig. 24).[154] In this method, the fluorotyrosines are temporarily masked from the cellular metabolism by using a light removable protecting group.[155] Like in the above mentioned studies, the single substitution with 3F-Tyr analogues results in a red shift, while the *ortho*-substituted analogue (2F-Tyr) results in a blue shift of fluorescence emission. This proves the ability of a single fluorotyrosine to change the electronic properties of the chromophore and essentially affect protein function.

Conversely, the global replacement of Tyr with 3-FTyr in DsRed-Monomer, a red fluorescent protein isolated from *Discosoma coral*, in which a Tyr residue is part of the fluorophore, resulted in a 12 nm blue shift in fluorescence emission.[156] The substitutions did not appear to have any effect on the overall folding pattern of the protein and the mutant displayed an improvement in quantum yield and fluorescence emission intensities.

3.3 Tryptophan

3.3.1 General properties. Fluorination at different positions of the indole ring of Trp (Fig. 25) results in inverted polarities and leads to an increased charge separation. The net charge of the substituted carbon atoms are reduced in the following order: 4-FTrp > 6-FTrp > 5-FTrp.[157] Hydrophobicities derived from 1-octanol/water partitioning experiments of Trp, 4-FTrp, 5-FTrp and 6-FTrp performed by two different groups, however, yielded contradictory results. Xu *et al.*, who undertook their studies with aqueous solutions of the amino acids, established the order of increasing hydrophobicity as: Trp < 4-FTrp < 6-FTrp < 5-FTrp.[157] Wong and Eftik, on the other hand, who repeated these measurements using pH 7.05 phosphate buffer instead of water, found an inversion in the order of 5-FTrp and 6-FTrp, indicating 6-FTrp as the most hydrophobic compound in this group.[158] These contradictory results stress the point that hydrophobicity of amino acids should be evaluated under conditions that are relevant to folding and activity studies especially when the side chains incorporate ionisable functional groups. Unfortunately, the hydrophobicities of Trp analogues with more than one fluorine substitution are not available.

Examination of the reaction kinetics of *E.coli* tryptophan indole lyase, which catalyses the reversible hydrolytic cleavage of L-tryptophan to indole, and ammonium pyruvate, revealed that 4,7-, 5,7-, and 7-FTrp are slower substrates for degradation than Trp.[159,160] Moreover, the intracellular uptake of 6-FTrp and 4-FTrp leads to death of breast cancer cells with antitumor efficacies that are similar to known chemotherapeutics.[161] This might be due to F-Trp incorporation into proteins and intracellular accumulation.

It is a popular practice to use biosynthetically incorporated fluorotryptophan analogues as 'intrinsic' probes of protein

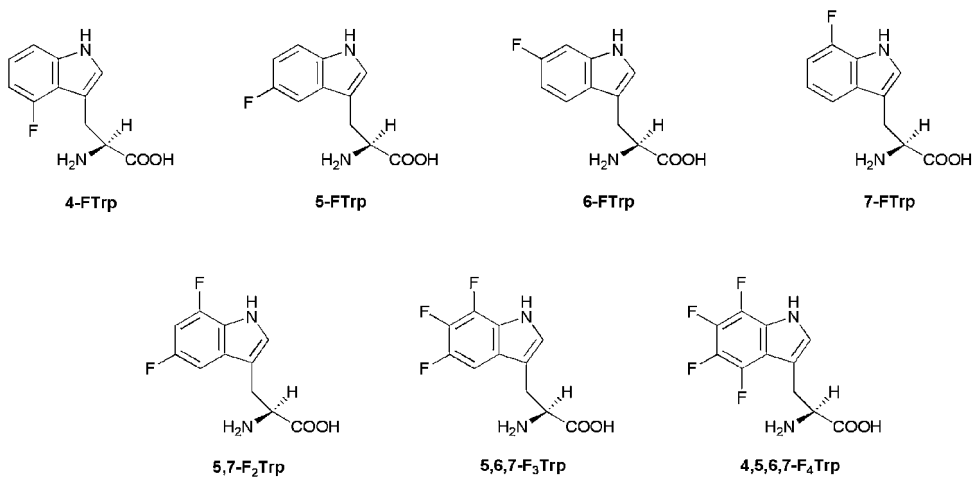


Fig. 25 Structures of available fluorinated tryptophan analogues.

local structures. The absorbance spectra of 5-FTrp extend to longer wavelengths than those of its natural counterpart and, thus, enable preferential excitation of the fluorinated analogue in fluorescence experiments. The absorbance of 4-FTrp is blue shifted and has a low fluorescence quantum yield at 285 nm (Trp : 4-FTrp quantum yield ratio > 100 : 1). This makes the analogue useful for studies in which it is desirable to greatly diminish the fluorescence of a particular Trp site.[162] For example, 4-FTrp as a non-fluorescent analogue allows identifying spectroscopic contributions of 'hidden' chromophores such as Tyr or Phe in proteins.[163]

3.3.2 Single substitutions. A protein with only one target residue can be used to study a single atom exchange, when substituting a single side chain hydrogen atom. In this respect, denaturation studies as well as phosphorescence and optically detected magnetic resonance (ODMR) measurements were carried out on fluorotryptophan analogues of staphylococcal nuclease.[158,164] 4-FTrp, 5-FTrp and 6-FTrp have been incorporated into the wild-type enzyme (which has a single Trp residue at position 140), the V66W mutant (with two Trp residues) and a fragment of the mutant (with the Trp residue at position 66) *via* SPI. While 5-FTrp and 6-FTrp did not alter the stability of the three investigated proteins, 4-FTrp seemed to have a modest stabilizing effect.[158] The phosphorescence spectra of all analogues were slightly red shifted and, although lacking a crystal structure, the obtained phosphorescence data combined with ODMR provided preliminary evidence that the structural integrity of the enzyme is retained when any of the investigated fluorinated Trp analogues were incorporated.[164]

Annexin V is a protein similar to staphylococcal nuclease in terms of the overall size and thermodynamic stability.[165] Applying SPI, the only Trp residue of human annexin V was also substituted with 4-FTrp, 5-FTrp and 6-FTrp.[166] In this study, the denaturation profile of the 5F-Trp mutant showed a slight increase in melting temperature, while the two other analogues exhibited similarly lowered T_m values when compared to the wild type. The Far-UV CD spectra of wild-type annexin V and its related mutants are almost superimposable, and the crystal structure of the 5- and 6-FTrp atomic mutants confirmed only minimal changes and a high degree of isomorphism with the wild type. Modelling data furthermore indicated a much tighter packing of the 5-FTrp into the local protein environment than 4-FTrp and 6-FTrp. 5-FTrp might thus cause less unfavourable electrostatic or steric interactions than its analogues, which in turn might give an explanation for the observed differences in the thermodynamic stability of these annexin V mutants. The Trp residue of annexin V is crucial for Ca^{2+}-dependent membrane binding and channel activity of the protein. Measuring the calcium influx through a liposome membrane revealed that 6-FTrp annexin V is three times more active than the wild type protein, while 4-Trp annexin V showed 32%, and the 5-Ftrp mutant only 50% of the wild type protein activity.

Applying nonsense codon suppression procedures and oocyte expression, 4-FTrp, 5-FTrp, 6-FTrp, 5,7-F_2Trp, 5,6,7-F_3Trp and 4,5,6,7-F_4Trp were incorporated at four different Trp sites of the mouse muscle nicotinic acetylcholine receptor (nAChR).[106] All four Trp sites appear to define the agonist binding site of this ligand gated ion channel, as they

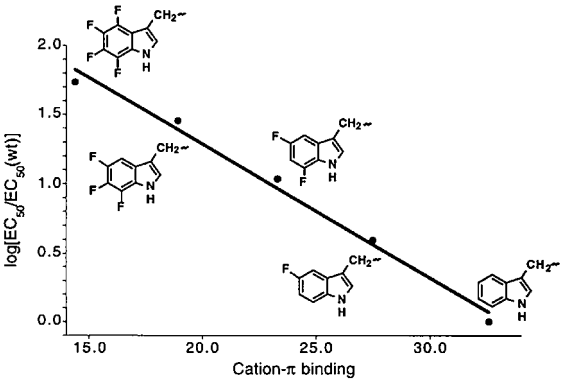


Fig. 26 Plot of log[EC_{50}/EC_{50}(wild type)] *vs.* quantitative cation–π binding ability at α149. The data fit the line $y = 3.2 - 0.096x$, with a correlation coefficient $r = 0.99$. Error bars are approximately the size of the markers. (reproduced with permission from Zhong *et al.*[106] Copyright 2011, National Academy of Sciences, USA).

may be involved in cation–π interactions that facilitate the binding of the ammonium group of the acetylcholine agonist.[167] The results revealed that only one of the four investigated substitution positions (α149) exhibits a correlation between fluorination and its resulting weakening of cation–π interaction and agonist binding (Fig. 26). Here, progressive fluorination leads to a systematic increase of EC_{50}.

3.3.3 Global substitutions. *E. coli* arginyl-tRNA synthase (ArgRS) containing 4-FTrp was biosynthetically prepared and purified from a tryptophan auxotrophic bacterial strain.[168] ArgRS is a 577 amino acid residue single peptide enzyme that possesses 5 Trp residues. Studies on the effect of substitution revealed that the first order rate constant is by approximately 20% lower than that of its native analogue but CD studies did not indicate any differences in the secondary structure within the limits of this analytical method. Further investigations on thermal unfolding behaviour showed that fluorination did not significantly affect thermal stability, even though the melting temperature of the fluorinated analogue was slightly lower. Furthermore, a 2 nm blue shift was observed in the UV spectrum of 4-FTrp ArgRS and the intrinsic fluorescence has a very low intensity which is presumably caused by traces of native ArgRS.

Furthermore, 4-FTrp, 5-FTrp, 6-FTrp and 7-FTrp have been globally incorporated into enhanced cyan fluorescence protein (ECFP) using SPI.[169] ECFP contains two Trp residues, one inside and one outside of the protein chromophore. All fluoro-Trp containing ECFPs exhibit a blue shift in emission and absorption maxima with the exception of 6-FTrp ECFP, which remains more or less unchanged. Surprisingly, the incorporation of the 'silent fluorophore' 4-Ftrp did not suppress the fluorescence of the protein. Moreover, measured pK_a values of the fluorinated analogues were only slightly reduced, while the wild type pK_a lay within the error range of the stated values. The crystal structure of 4-FTrp-ECFP shows the fluoro-indole moiety of the chromophore tightly packed with the fluorine atom in the range of hydrophobic contact (3.5 Å) with the sulfur atom of a Met residue. Systematic mutagenesis experiments showed that this Met residue, unlike

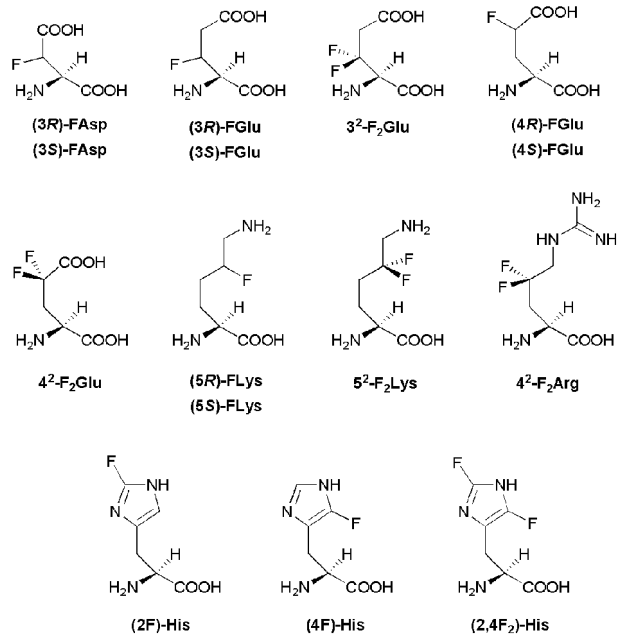


Fig. 27 Structures of side-chain fluorinated charged α-L-amino acids for which syntheses have been reported.

the four other Met residues of the protein, cannot be substituted with other canonical amino acids without a loss of fluorescence and protein folding, thus confirming a possible sulfur–aromatic interaction that might be contributing to the overall structural stability of the protein.

More recently, molecular dynamic simulations on wild type and modified streptavidin (SAV) were performed to estimate the impact of tryptophan fluorination (4-FTrp, 5-FTrp, and 6-FTrp) on biotin binding ability.[170] The high affinity of the non-covalent biotin binding to streptavidin relies on multiple H-bonding interactions combined with a deep hydrophobic pocket primarily formed by four tryptophan residues.[171] Computational investigations predicted protein stability and folding to be marginally affected by fluorination, but the 4-FTrp-SAV binding pocket seemed to reduce its ability to form numerous hydrogen bonds. Meanwhile 5-FTrp-SAV was predicated to form a more stable ligand–host hydrogen bonding network in comparison to the wild type protein.[170] These results indicate that 5-FTrp-SAV may be promising for streptavidin modifications. However, these predictions still must be verified by experimental data.

4. Analogues of polar and charged amino acids

4.1 General properties

Trifunctional amino acids such as those carrying charged and polar side chains engage in tertiary interactions that stabilize the protein structure and substrate binding as well as mediate catalysis in enzymatic reactions. Fluorine exerts a strong electronic effect on the properties of charged and polar functional groups.[17] It decreases the pK_a of carboxyl side chains (Asp, Glu) thereby stabilizing their negative charge while destabilizing the positive charge of protonated ammonium (Lys), imidazolium (His) or guanidinium moieties (Arg).

Side chains incorporating functional groups such as hydroxyl (Ser, Thr) and amide groups (Asn, Gln) often engage in hydrogen bonds and may temporarily form covalent bonds to substrates over the course of enzymatic reactions. Their side chains furthermore serve as sites for posttranslational modifications such as phosphorylation and glycosylation. Fluorine's inductive effect may also strongly affect these interactions. Although the alteration of acid–base properties and nucleophilicity/electrophilicity may be useful to modulate both protein stability and activity, fluorinated analogues of polar and charged amino acids have rarely been used to modify peptides and proteins. This may be due to difficulties that arise during syntheses either from instability of the final unprotected products or their fluorinated precursors. Fluorinated analogues of protected serine and cysteine have been synthesized.[172] However, the free analogues are unstable and peptides containing these residues have, to the best of our knowledge, never been reported. Fig. 27 and 28 summarize the derivatives of charged and polar amino acids that are synthetically available. Of all these residues only Glu, Asn and Thr derivatives have been incorporated into peptides and a few examples of proteins containing fluorinated His analogues have been reported.

4.2 Glutamic acid and histidine

(3*S*)- and (3*R*)-FGlu have been used to probe mechanistic details of Vitamin K-dependent carboxylation at the γ-carbon

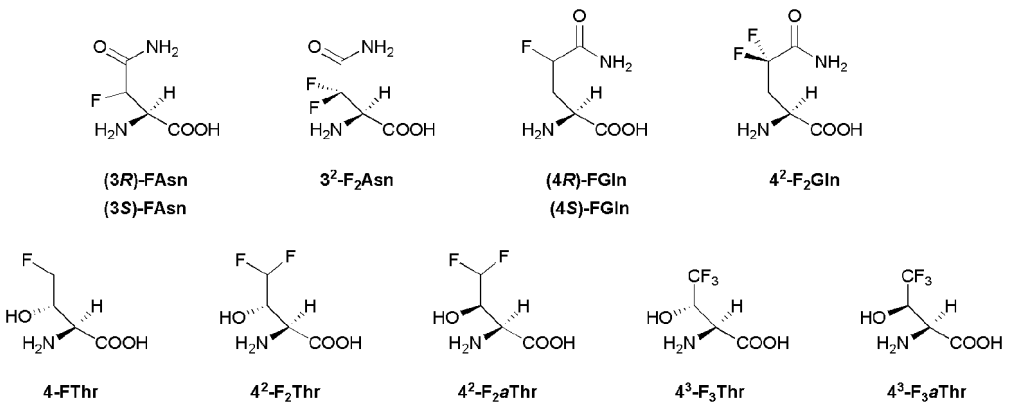


Fig. 28 Structures of side-chain fluorinated polar α-L-amino acids for which syntheses have been reported (*a* = *allo*).

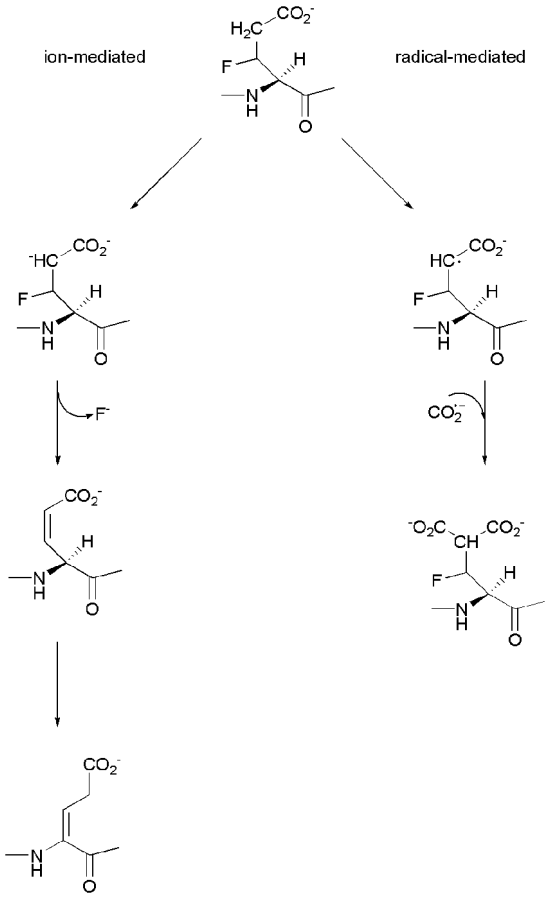


Fig. 29 Schematic representation of the two possible mechanisms. A radical mechanism would not lead to elimination of a fluoride radical due to the strength of the C–F bond. Thus, formation of dehydroglutamate indicates formation of an intermediate carbanion (according to Vidal-Cros *et al.*)[173]

of glutamic acid using the tripeptide (Boc-Xaa-Glu-[^{3}H]Val).[173] The authors hypothesized that since β-fluoro carbanions are prone to fluoride elimination, they should be able to distinguish a radical-mediated mechanism from ion-mediated mechanism by identifying the final product (Fig. 29). Indeed, they were able to show that dehydroglutamate is formed by elimination of fluoride from an intermediate carbanion. Importantly, only the peptide containing (3*S*)-FGlu eliminated fluoride, and the peptide containing (3*R*)-FGlu did not convert. These findings could be explained by the unfavourable relative orientations of H4 and F3 for the elimination reaction.

Hartman *et al.* found that both 4-fluorinated Glu analogues as well as 4,4-F_2Glu are good substrates for enzymatic amino-acylation of tRNA and that they are efficiently incorporated into model peptides.[174,175] However, to our knowledge no peptide has been systematically investigated for the effects of 4-FGlu analogues on their structure and activity. In fact, fluorinated analogues of Glu may actually be useful to probe the role of catalytically active Glu residues.

Histidine's imidazolyl side chain is crucial to many protein functions. These functions include the activity as a proton acceptor and donor in a general acid/base catalysis and coordination of metal ions that may act as cofactors in catalysis or protein structure stabilization. Incorporation of fluorine atoms into position 2 or 4 of the imidazole ring lowers the pK_a values of the imidazolium cation by several pK_a units from 6.0 to 1.22 for (2F)-His and to 1.76 for (4F)-His.[176] The corresponding pK_a value for (2,4F_2)-His has not been reported. These dramatic changes in pK_a provide the means to assess the role of histidine in pH-dependent biological processes.

Jackson *et al.* used (4F)-His to study the role of two catalytic His residues within ribonulcease A. Three variants in which His12, His119 or both were replaced by (4F)-His have been synthesized.[177] Although large changes in the pH–rate profile for the cleavage of uridyl-3′,5′-adenosine were observed, the relative catalytic rates were only slightly reduced. These results confirmed that His12 and His119 act as general acid/base catalysts during the reaction. However, the proton transfer steps are not rate limiting to the cleavage reaction. Thus, fluorinated histidines were shown to be formidable tools to probe enzyme catalysis involving His as the acid/base catalyst.

(2F)-His has furthermore been used to label the protective antigen (PA) component of the anthrax toxin.[178,179] Generally, binding of PA to the host cell receptor will lead to the proteolytic cleavage of PA and formation of a heptameric pore in cell membrane, which enables edema factors and lethal factors to enter the host cell. Pore formation depends on a conformational change in the heptameric prepore that occurs in the neutral to mildly acidic pH range. It is assumed that protonation of one or more amino acids triggers this conformational change and eventually leads to the pore formation. Because the pK_a value of (2F)-His is much lower than that of histidine, Wimalasena *et al.* employed this analogue to study the pH-dependent mechanism of pore formation. They found that, although the lower pK_a of (2F)-His is expected to hinder protonation, labelled PA undergoes the conformational transition within the same pH-range as the wild type. Thus, His protonation does not trigger pore formation. However, the presence of (2F)-His within PA prevented translocation of the lethal factor into the cytosol. In their follow-up study the authors identified that fluorination of His299 caused the defect in translocation.[128] These more recent data could support two separate models of pore formation: (1) when PA is bound to the cellular receptor; the mechanism of pore formation is supported by protonation of histidine and (2) when the receptor is not bound protonation of histidine does not mediate pore formation.

Both (2F)-His and (4F)-His) have also been successfully incorporated into the chaperone PapD by means of expression in *E. coli* strains auxotrophic for histidine.[180] However, functional and structural studies with the mutant proteins have not been reported.

4.3 Asparagine and threonine

Among the fluorinated analogues of polar amino acids only (3*S*)-FAsn and 4^3-F_3Thr have been used to modify peptides. (3*S*)-FAsn and its enantiomer have been incorporated into the short synthetic peptide Ac-**Asn**LeuThr-CONHMe, upon which it failed to undergo enzymatic glycosylation in a cell-free system even at very high concentration. Together with the findings from *in vivo* studies these results supported the idea

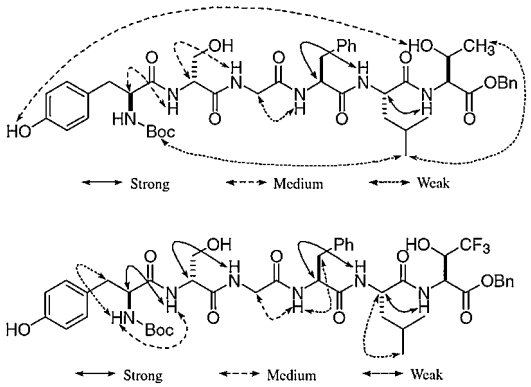


Fig. 30 Schematic representation of the fluorination effect on the conformation of a hexapeptide by introduction of fluorinated analogue of threonine. Arrows represent observed NOEs (reproduced with permission from Kitamoto *et al.*[184] Copyright 2011, The Chemical Society of Japan).

that incorporation of 3-FAsn into cellular protein exerts its toxicity by inhibition of *N*-linked glycosylation.[181–183]

Kitamo *et al.* reported the replacement of threonine with 4^3-F_3Thr in an enkephalin-derived hexapeptide, Tyr-D-Ser-Gly-Phe-Leu-Thr (DSLET) to investigate conformational effects induced by replacing threonines CH_3 group by a CF_3-group.[184,185] The natural peptide DSLET showed that the estimated helix or spherical structure is stabilized by two characteristic long-range interactions (hydrogen bond) between the amino acids five or six residues apart from each other. The introduction of a trifluoromethyl group led to loss of the original structure, as shown by the disappearance of these long-range hydrogen bonds. This example suggests that a significant conformational change can be induced by the strong electron withdrawing effect of fluorine within the side chain of Thr (Fig. 30).

Although it has never been applied in peptide and protein engineering, it may be worth mentioning that the monofluorinated analogue 4-FThr is the only naturally occurring fluorinated amino acid.

5. Analogues of prolin

5.1 General properties

Prolyl peptide bonds are structurally unique since Pro is a cyclic amino acid with a side chain incorporating both, N^{α} and C^{α}. In a polypeptide the cyclic side chain imposes steric constraints onto the backbone that lead to peculiar structural features, which usually make it incompatible with regular secondary structures such as α-helices and β-sheets.[186,187] As a consequence of inducing a kink in peptide chains, proline is often found in loop regions or turn structures. Furthermore, polyprolines adopt unique helical structures such as polyproline I (PPI) and polyproline II (PPII) helices.[188] Fluorinated analogues of proline that have been synthesized are shown in Fig. 31.

The prevalent conformation of the pyrrolidine ring (C^{γ}-*endo vs.* C^{γ}-*exo*, Fig. 32) that may be affected by substituents can have a pronounced impact on the thermodynamics and kinetics of folding. A more profound discussion of fluorine's impact in this regard is required to understand its effects in the context of protein structure and protein–protein interactions.

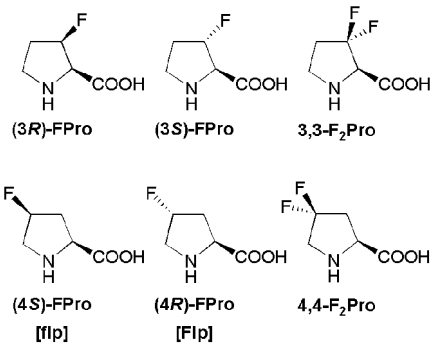


Fig. 31 Structures of side-chain fluorinated L-proline derivatives that have been incorporated into peptides and proteins. Letter codes that are often used in the literature are given in brackets.

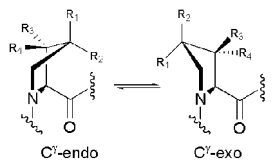


Fig. 32 Equilibrium between the two most abundant ring conformations of Pro (R_1–R_4 are specified in Table 1).

Conformational properties of molecules incorporating fluorine substituents at C–C single bonds can often be rationalized on the basis of the *gauche* effect.[11] Somewhat counterintuitively at first sight, fluorine substituents and another electronegative substituent attached to the adjacent carbon adopt a *gauche* rather than an *anti* orientation, which could be expected based on steric arguments (Fig. 33). This effect is especially pronounced when the other substituent is an amide[189] as in the case of C3- or C4-substituted Pro within in a polypeptide chain. The strong *gauche* preference of the β-fluoro amide moiety has an important consequence on the conformation of the pyrrolidine ring as seen in Fig. 33. Depending on the stereochemistry of the substituted carbon, the ring is found either in the C^{γ}-*endo* or the C^{γ}-*exo* conformation, with a preference for a *gauche* orientation of the amide and the fluorine substituent.

Taking the *gauche* effect as the only criterion, the preference for the C^{γ}-*endo* or C^{γ}-*exo* conformer as summarised in Table 1 can readily be predicted for (4*R*)-, (4*S*)-, (3*R*)-, (3*S*)-FPro as well as Hyp and hyp. Because the pyrrolidine ring is unsubstituted in native Pro, the accompanying *gauche* effects are absent. Thus, although the acylated methyl ester crystallizes with a C^{γ}-*endo* pucker,[190] its actual equilibrium population in

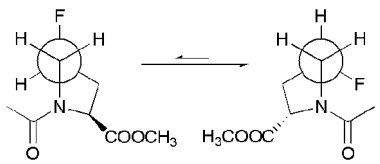


Fig. 33 Idealized Newman projection of the C4–C5 single bond of Ac–(4*S*)-FPro–OMe showing its preference for the *gauche* conformer, in which the pyrrolidine ring adopts the C^{γ}-*endo* conformation.

Table 1 Conformational preferences of the Pro pyrrolidine ring in Ac–Pro–OMe derivatives. The same parameters are given for Hyp and its stereoisomer hyp for comparison, since Hyp frequently occurs in native proteins. Values for $K_{trans/cis}$ refer to aqueous solution

Amino Acid	R_1	R_2	R_3	R_4	C^γ-*endo*	C^γ-*exo*	$K_{trans/cis}$ (25 °C)	Ref.
Pro	H	H	H	H	x		4.6(4.8)[b]	190, 191,[a] 192, 193
(4*R*)-FPro [Flp]	**F**	H	H	H		x	6.7(6.9)[b]	190, 192–194
(4*S*)-FPro [flp]	H	**F**	H	H	x		2.5(2.5)[b]	191–194
(3*R*)-FPro	H	H	**F**	H		x	8.9	195, 196
(3*S*)-FPro	H	H	H	**F**	x		4.3	195, 196
4^2-F_2Pro	**F**	**F**	H	H	x		3.4	193[b]
3^2-F_2Pro	H	H	**F**	**F**	n.d[c]	n.d[c]	n.d[c]	
Hyp (4*R*)	**OH**	H	H	H		x	6.1	190, 192
Hyp (4*S*)	H	**OH**	H	H	x		2.4	192

[a] In solution Ac–Pro–OMe shows only a slight preference for the *endo* conformation (66% population). [b] The $K_{trans/cis}$ values in parentheses have been calculated from ref. 193 using the equation describing its temperature dependence and taking 298.15 K as the temperature and 8.3145 J mol^{-1} K^{-1} as the gas constant *R*. [c] 3^2-F_2Pro has been synthesized[197] but never characterized in this respect.

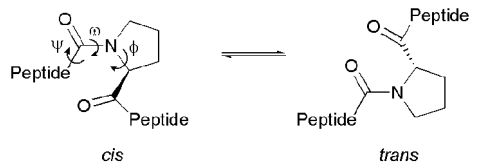


Fig. 34 *Cis*/*trans* equilibrium of prolyl peptide bonds.

dioxane solution is only 66%.[191] Likewise, disubstitution of C4 leads to counteracting *gauche* effects of the vicinal fluorine atoms making 4,4-F_2Pro very similar to native Pro.

Due to its cyclic structure the conformation of the prolyl side chain directly affects the main-chain torsional angles ϕ, ψ and ω (Fig. 34). Therefore, the *cis*/*trans* equilibrium of Xaa-Pro peptide bonds, at least in part, depends on the C^γ-pucker.[191] Steric considerations suggest a strong *trans* preference of the peptide bond. However, *cis* peptide bonds do occur with a much higher probability when Xaa-Pro units are involved (5.2% compared to 0.03% for Xaa-nonPro peptide bonds in peptides and proteins of known structure).[198] Crystal structure analyses of different sets of peptides and proteins revealed that *cis*-Xaa-Pro peptide bonds are mostly associated with a C^γ-*endo* pucker while, analyses of *trans*-Xaa-Pro peptide bonds did not reveal such a preference.[199]

Computational studies as well as NMR investigations[191] support the conclusion that the *exo* pucker preorganises the backbone in a way that allows for favourable $O_{i-1} \rightarrow C_i{=}O_i$ $n \rightarrow \pi^*$ interactions between the amide oxygen and the prolyl carboxyl group. It is has been found that such $n \rightarrow \pi^*$ interactions generally stabilise *trans* over *cis* peptide bonds in addition to preventing steric strain. Close contacts between the amide oxygen and the Pro carboxyl carbon were found in crystals of Ac–(4*R*)-FPro–OMe[196] and Ac–(3*S*)-FPro–OMe.[190] Furthermore, infrared spectroscopy[192] and computational chemistry[191] verify the presence of these interactions. Accordingly, fluorine substituents that favour the C^γ-*exo* pucker, *i.e.* (4*R*)- and (3*R*)-FPro, have a comparably high equilibrium constant for the *trans* to *cis* isomerization while those that favour the C^γ-*endo* pucker exhibit a smaller preference for the *trans* peptide bond. It may be important to note that the C^γ-*endo* conformation does not lead to an actual preference of *cis* over *trans*, since all $K_{trans/cis}$ values are greater than one (Table 1). It does, however, imply that the energy difference between *cis* and *trans* is smaller[190] for the C^γ-*endo* pucker and thus the probability to form a *cis* peptide bond is higher for Pro, (4*S*)- and (3*S*)-FPro as well as 4^2-F_2Pro.

The strong electron withdrawing effect of fluorine as evidenced by a distinct decrease in pK_a of proline's NH group (see Table 2) has another very important effect on the kinetics of the *cis*/*trans* isomerization. Vibrational spectroscopy indicates that the bond order of the amide C=O of substituted prolines increases along with a decreasing pK_a.[193,200] At the same time the bond order of the amide bond decreases and more closely resembles a single rather than a double bond. A rotation along a single bond-like peptide linkage, therefore, is facilitated. Accordingly, the activation barrier for isomerization decreases in the order Pro ≈ Hyp > (4*R*)-FPro > (4*S*)-FPro > 4^2-F_2Pro.

Comparable investigations on C3 fluorine substituted Pro have not been published. Table 2 contains predicted pK_a values for these derivatives indicating that they are slightly lower for the monofluorinated C3 derivatives and significantly lower for 3^2-F_2Pro. However, a prediction of the effect on isomerization kinetics would be speculative and remains an aspect of further investigation. The finding that fluorine's inductive effect may be useful to modulate *cis*/*trans* isomerization kinetics is of great importance since prolyl peptide bond isomerization is a rate limiting step in protein folding.

Table 2 pK_a values for the NH of the free amino acid and rotational barriers at 300 K for the *trans* → *cis* and *cis* → *trans* conversion determined using Ac–Pro–OMe derivatives

Amino acid	pK_a	$\Delta G^{\neq}$/kJ mol^{-1} *trans* → *cis*	$\Delta G^{\neq}$/kJ mol^{-1} *cis* → *trans*	Ref.
Pro	10.8	88.4	84.5	193, 200
(4*R*)-FPro [Flp]	9.23	87.4	82.6	193, 200
(4*S*)-FPro [flp]	(9.29)[a]	86.3	84.0	193
4^2-F_2Pro	7.15	83.8	80.8	193
(3*R*)-FPro	(8.37)[a]	n.d.	n.d.	
(3*S*)-FPro	(8.37)[a]	n.d.	n.d.	
3^2-F_2Pro	(6.05)[a]	n.d.	n.d.	
Hyp	9.68	88.2[b]	87.7[b]	200
Hyp	(10.3)[a]	n.d.	n.d.	

[a] Predicted pK_a calculated using Advanced Chemistry Development (ACD/Labs) Software V11.02 (1994–2010 ACD/Labs). [b] Calculated using the Gibbs–Helmholtz equation using the activation parameters reported in ref. 200 taking 300 K as the temperature to allow direct comparison.

At this point it may be worth summarising the observations reviewed above:

• Fluorine at C3 and C4 favours the C^{γ}-*exo* pucker when substituted for the *proR* hydrogen; the opposite is true for substitution of the *proS* hydrogen.

• The C^{γ}-*exo* pucker favours a conformation in which an $O_{i-1} \rightarrow C_i{=}O_i$ n $\rightarrow \pi^*$ interaction between the amide oxygen and the prolyl carboxyl group allows for additional stabilization of the *trans*-peptide bond; the energy difference between *cis* and *trans* is reduced for the C^{γ}-*endo* pucker, *i.e.* adopting a *cis*-peptide bond is more likely.

• Fluorination accelerates the *cis*/*trans* isomerization. This effect correlates with the pK_a of the amino group of substituted prolines.

5.2 Single substitutions

5.2.1 Effects on enzymatic reactions and interactions with enzymes. As will be discussed in more detail in Section 5.3, hydroxylation at C4 of the prolyl side chain is an essential posttranslational modification. Several model peptides containing fluorinated prolines have been tested as substrates/inhibitors for human prolyl 4-hydroxylase (P4H) which catalyses the hydroxylation at C4 of Pro in procollagen strands. In an attempt to gain mechanistic insights, the tetrapeptide Z-Gly-Phe**Pro**Gly-OEt was used as a model by Tandon *et al.*[201] Several analogues have been synthesized that, among other modifications, contained 3-FPro. Interestingly, fluorination was achieved by directly fluorinating a 3-(OH)Pro containing tetrapeptide using DAST. However, the fluorination reaction was not stereospecific and a separation of the diastereomeric peptide mixture has not been pursued. It was shown that the peptide containing diastereomeric 3-FPro is readily hydroxylated at C4 by P4H but reaction rates have not been reported. It would be worth investigating (3*S*)-FPro and (3*R*)-FPro individually since both diastereomers exhibit distinct conformational preferences that may have an impact on the kinetics. This is interesting because prolyl conformation as well as the *trans*/*cis* ratio of the Xaa-Pro peptide bond, apparently, may have a strong impact on substrate binding as has been shown for 4-FPro diastereomers by Gorres *et al.*[202] In their study a similar peptide, PEG-GlyTyr**Pro**Gly-OEt, was used as a model. The peptide containing (4*R*)-FPro was not converted because P4H abstracts the *proR* hydrogen from C4 and this position is blocked by fluorine. Moreover, it did not inhibit the enzyme showing that it does not bind to the active site. In contrast, the peptide containing (4*S*)-FPro, though at a reduced rate, was readily hydroxylated. However, due to the instability of the initially formed α-fluorohydrin spontaneous hydrogen fluoride elimination released 4-ketoproline as the final product. These results strongly support the conclusion that binding and conversion must strongly depend on conformational aspects of the Xaa–Pro peptide bond. Pro and (4*S*)-FPro prefer the C^{γ}-*endo* pucker and comparatively low *trans*/*cis* ratios while (4*R*)-FPro as well as Hyp prefer the C^{γ}-*exo* pucker and high *trans*/*cis* ratios. Peptides containing these latter residues are neither converted by P4H nor do they inhibit the enzyme. Thus, by using fluoroprolines as probes, some important conclusions could be made concerning substrate specificity of P4H. Substrates similar to Hyp with regard to conformational preferences are not recognized by the enzyme. It may be brought to the reader's attention that these more recent studies contradict reports from the 1960s in which (4*R*)-Pro was found to be hydroxylated.[203,204] However, these earlier studies relied on more indirect assays the results of which might have been mistaken. Nevertheless, the results gained with (4*S*)-FPro inspired the development of the first direct and continuous assay of P4H based on fluoride detection.[205] Fluorinated proline as a probe may, therefore, facilitate high throughput screenings for novel inhibitors in future applications.

Another important class of enzymes addressing prolyl peptide bonds are peptidyl prolyl *cis*/*trans* isomerases (PPIases). These enzymes are crucial for protein folding since the *cis*/*trans* isomerization is a rate limiting step in structure formation. The catalytic efficiency of four different human as well as bacterial PPIases to promote the *cis* $\rightarrow$ *trans* isomerization was tested by Golbik *et al.*[206] using the model peptide succinyl-AlaSer**Pro**Phe-*p*-nitroanilide. As outlined in Section 5.1, fluorination at C4 generally accelerates the spontaneous isomerization reaction such that in the case of 4,4-F_2Pro the non-catalyzed reaction of the model peptide was already too fast to enable reliable measurements with the enzymes. The enzymes were mostly less efficient in catalyzing the isomerization of the fluorinated model peptides. However, for most of them the (4*S*)-fluorine substitution conferred higher catalytic susceptibility to the substrates than did (4*R*)-FPro. Importantly so, this trend was reversed for the enzyme *E. coli* Par10. In line with this study, comparable investigations have been carried out with a bigger protein, the barstar mutant (Cys40Ala/Cys82Ala/Pro27Ala)[193] containing only one Pro at position 48 that engages in a *cis*-TyrPro peptide bond in the native state. Five different enzymes that promote refolding of the protein by catalyzing the *trans* $\rightarrow$ *cis* isomerization were tested against analogues containing both 4-FPro diastereomers as well as 4,4-F_2Pro. Although the reverse reaction was investigated with barstar, the trend that (4*S*)-FPro confers higher enzymatic susceptibility than (4*R*)-FPro was shown to persist and, again, a reverse trend was observed for *E. coli* Par10. The stereochemical discrimination, however, was less powerful. Moreover, with the exception of the enzyme *E. coli* TF that showed an activity of 2250%, all other enzymes were practically inefficient in accelerating the isomerization of the $Tyr4^2$-F_2Pro peptide bond. However, in contrast to the trend found for the model peptide, some of the enzymes were significantly more efficient in the context of the fluorinated barstar analogues compared to the wild type protein. This finding, apparently, is sequence specific and may depend on the residues flanking the proline derivative. The authors furthermore noted that the zero susceptibility of 4,4-F_2Pro containing barstar towards most PPIases tested may be explained by invoking fluorine's strong inductive effect that renders the peptide linkage more single bond like. A single bond should more closely resemble the transition state of the isomerization. Therefore, the energy difference between the ground and transition state binding would be reduced for 4,4-F_2Pro containing peptides. Accordingly, transition state binding would be disfavoured and so would be catalysis.

Comparable considerations might also apply to proteolytic enzymes that specifically cleave Xaa–Pro peptide bonds. Surprisingly, detailed investigations covering the proteolytic stability of peptides containing fluoroprolines are very rare. Phe–Pro peptide bonds, for example, are of interest for the development of inhibitors against HIV protease since, in contrast to mammalian endopeptidases, it specifically cleaves Tyr/Phe–Pro peptide bonds at the N-terminus of Pro. Guedj and co-workers synthesised a diastereomeric mixture of the hexapeptide Boc-AlaAlaPhe**4-FPro**ValVal-OMe by direct fluorination of the 4-hydroxyproline containing precursor.[207] Unfortunately, the peptide was found to be inactive against HIV protease. The authors, however, did not mention whether the lack of inhibitory activity is due to degradation of the peptide or whether it does not bind to the enzyme.

5.2.2 Effects of fluorination on folding and stability

Collagen and other polyproline models. Collagen is the most abundant protein in animals. It is a major component in connective tissue where it forms fibrils composed of tightly wound triple PPII helices that are extraordinarily stable. Its primary structure is often based on repetitions of Pro-rich sequences such as ProHypGly or GlyProHyp. Most studies covering the impact of fluorination on the collagen structure and stability were conducted on models in which Pro was globally substituted by fluorinated analogues. These studies will be discussed in detail in Section 5.3. Fewer studies were published dealing with single substitutions within so-called host–guest models. Malkar *et al*. substituted a single (4*R*)-FPro or (4*S*)-FPro for Hyp (substituted position in bold) within the α1-region (*italic*) of a type IV collagen host sequence (GlyPro-Hyp)$_3$-GlyPro**Pro**-*GlyValLys-GlyAspLys-GlyAsnPro-GlyTrpPro-GlyAlaPro*-(GlyProHyp)$_4$-NH_2.[208] This region in collagen is known to promote melanoma cell adhesion, spreading and signalling. They found that while (4*R*)-FPro stabilises the triple helix, its diastereomer (4*S*)-FPro slightly destabilises it compared to Hyp, and yet it forms a triple helix. These results may be explained by the fact that (4*R*)-FPro more strongly exhibits a conformational bias that equals that of Hyp thereby favourably preorganizing the backbone. In contrast, (4*S*)-FPro favours an opposite conformation and thus hampers folding. They furthermore found that such a single substitution can have a measurable effect on melanoma cell adhesion and spreading. Briefly, the stability of the triple helix directly correlates with both, spreading and signalling, although the authors state that these effects might not primarily arise from FPro substitutions *per se*, but rather from its overall effect on the equilibrium between unfolded, folded and partially folded states of the triple helix. Interestingly, Persikov *et al*. found a contrary stability trend for Hyp and (4*R*)-FPro in the different triple helical host–guest model AcGly-(ProHypGly)$_3$-Pro**Pro**Gly-(ProHypGly)$_4$-Gly-NH_2.[209] Here, (4*R*)-FPro, although preferring the same conformation as Hyp, was found destabilizing compared to both, Pro and Hyp. (4*S*)-FPro, however, was not investigated. In a third model, Ac-(ProProGly)$_4$-Pro**Pro**Gly-(ProProGly)$_5$-NH_2, Shoulders *et al*. found a stabilizing effect of (4*R*)-FPro compared to Pro again.[210] It is obvious that the effects of single substitutions strongly depend on the immediate environment of the host sequence. Secondary effects on the residual monomer structure or hydration network come into play making it hard to explain eventual discrepancies between the effects of single substitutions in different models of the collagen repeat.

An important finding of the study on PPIase activity[206] was that the kinetic and thermodynamic trends of the *cis*/*trans* isomerization observed using fluorinated Ac–Pro–OMe models (see Section 5.1) are directly reflected in the spontaneous isomerization reactions of the model peptide as well as barstar. Also, Boulègue *et al*. investigated the folding kinetics of the N-terminal domain of minicollagen-1.[211] In its native state the Ala23-Pro24 peptide linkage is found in a *cis* conformation which explains the accumulation of a *trans* intermediate during oxidative folding due to the slow *trans* → *cis* isomerization. Incorporation of the strongly *trans* favouring (4*R*)-FPro led to a significantly enhanced folding intermediate that is only incompletely transformed into the final native fold. In contrast, the intermediate is formed to a much lesser extent in the case of (4*S*)-FPro and is almost completely transformed into the final *cis* product.

Fluorinated Pro in loop regions and turns. Both diastereomers of 4-FPro have been used to modify the insect and shellfish neurohormone proctolin (ArgTyrLeuProThr) to study their impact on peptide conformation by NMR.[212] Interestingly, the intrinsic conformational preferences of these fluorinated proline analogues do not seem to be decisive in the context of this peptide. In fact, the predicted conformations are exactly reversed, at least under the conditions applied for these experiments (all NMR investigations have been carried out in DMSO-d_6). The authors found that the intrinsic propensity of the peptide to form a putative γ-turn by establishing a $_{Leu}C{=}O\cdots HN_{Thr}$ hydrogen bond apparently overrides the conformational bias of (4*R*)-FPro and (4*S*)-FPro to adopt the C^{γ}-*exo* and C^{γ}-*endo* pucker, respectively. Moreover, the peptide containing (4*R*)-FPro adopts a *cis* Leu-Pro peptide bond rather than the predicted *trans* conformation. The authors explain these findings by invoking unfavourable steric interactions with the backbone or the Leu side chain, respectively, that would occur for theoretically favoured side chain conformers. It is important to mention that in a recent crystal structure of proctolin neither does the Pro side chain show a distinct ring pucker, nor does the crystal establish the postulated γ-turn.[213] Therefore, molecular modelling as well as crystal structures of the fluorinated proctolins would help to gain further conclusions. A brief biological testing showed that only the peptide containing (4*R*)-FPro is active against *tortrix*, while only the native peptide is active against *spider mite*. Detailed receptor binding studies were not pursued. Nevertheless, the structural studies indicate that intrinsic conformational properties evoked by fluorine substituents are not generally reflected in protein folds in which other interactions such as hydrogen bonds may be crucial to structure formation. While the recent analysis of proctolin crystals questions the formation of a γ-turn, it indicates some spatial proximity of the Pro side chain and the aryl moiety of Tyr residue. As will be detailed below, aromatic–prolyl CH$\cdots\pi$ interactions[214] may have to be considered when interpreting the results of proctolin modification.

Thomas *et al*. presented a systematic study in which they showed that such $^{Pro}CH\cdots\pi$ interactions may affect the *cis*/*trans* preference of substituted proline analogues using a

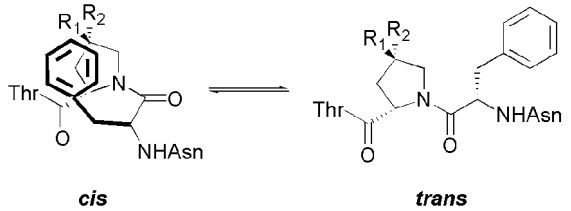


Fig. 35 *Cis*/*trans* equilibrium of the model peptide showing the interaction of the aromatic and prolyl residue in the *cis* conformation. Phe was chosen as an example. Nomenclature: (4*R*)-FPro (R_1 = F, R_2 = H), (4*S*)-FPro (R_1 = H, R_2 = F).

model peptide TXaaYaaN, where Xaa were different aromatic residues and Yaa were C4 substituted proline analogues (Fig. 35).[215]

They found that electron-rich aryl side chains such as Trp enhance the *cis* preference of (4*S*)-FPro, while electron-poor aryl side chains such as Tyr increase the bias for *trans* peptide bonds adopted by (4*R*)-FPro. In a *cis* conformation the peptide adopts a turn like structure that gains enthalpic stabilization by a favourable $_{Pro}{}^{\gamma}CH\cdots\pi$ interaction. This interaction is especially stable for Trp and (4*S*)-FPro since, in that case, the electron-rich Trp side chain directly interacts with a CH that is strongly polarised by the adjacent fluorine substituent (see Fig. 35). In conclusion, favourable or unfavourable aromatic–prolyl interactions, apparently, amplify conformational preferences of fluorine substituted Pro analogues. One may, however, hypothesize that depending on the geometry of such interactions, they could also have an opposite effect.

Zheng *et al.* studied another interaction of Pro and Trp in the context of the villin headpiece subdomain HP36 (Fig. 36).[216] In this structure Pro62 occupies a loop connecting the second and third α-helix and is involved in an aromatic–prolyl interaction with Trp64. In the NMR structure Pro is found C^{γ}-*exo* puckered although the displacement of C^{γ} from the ring plane is not very large and thus the pucker is not very pronounced. Most strikingly, (4*R*)-FPro that favours the C^{γ}-*exo* pucker was shown to destabilize HP36 in contrast to the C^{γ}-*endo* preferring (4*S*)-FPro. Thus, the authors concluded that the strong preorganization of the ring pucker by fluorination has an effect on the aromatic–proline interaction. Although electronegative fluorine substituents should render the $^{\gamma}$CH more electropositive, a strong C^{γ}-*exo* pucker would move it away from the ring, thus weakening the interaction. In contrast to other C^{γ}-*endo* preferring 4*S* substitutions such as OH, (4*S*)-FPro was the only substitution to slightly stabilise HP36. Therefore, the authors reasoned that in addition to moving the prolyl ring towards the aromatic ring by imposing a C^{γ}-*endo* pucker, burial of the hydrophobic C–F bond would also have a stabilizing effect.

In contrast, its exposure to the solvent in the case of (4*R*)-FPro would lead to destabilization. However, while their NMR studies indicated that the overall secondary structure is maintained for all variants, they did not directly address the issue of whether the conformational bias of the fluorinated prolines to prefer *exo* or *endo* puckers, respectively, actually persists in this context. Nevertheless, these studies add to the discussion that unfavourable interactions of a single C–F bond (with the solvent) may destabilize a protein. When side chain interactions involving the prolyl ring are absent, which is the case for solvent exposed residues, conformational biases often seem to be well reflected in the context of proteins as exemplified by investigations of the Trp cage protein (Fig. 37).[218] In this structure Pro12 defines a loop that connects the N-terminal α-helix and the C-terminal polyproline helix. It clearly exhibits a C^{γ}-*exo* pucker and is not involved in other interactions.

The authors assumed that C^{γ}-*exo* favouring residues such as (4*R*)-FPro would preorganize the backbone in a way that matches the native structure thus stabilizing the fold. Indeed, while (4*R*)-FPro that prefers the C^{γ}-*exo* pucker increases the overall helicity and stability of the protein, its diastereomer (4*S*)-FPro favours opposite conformations and, therefore, has a contrary effect. Interestingly, the other three Pro residues present in the structure of the trp cage miniprotein have not been investigated.

Detailed structure activity studies with peptides containing fluorinated prolines are rare. In one recent study Armishaw *et al.* investigated the antagonistic activity of two α-conotoxin analogues (IM1 and PnIA[A10L]) that target the α7 nicotinic acetylcholine receptor subtype (α7 nAChR).[220] A structural feature common to most α-conotoxins is a conserved Pro at position six that presumably induces a 3_{10}-helical turn in the backbone (Fig. 38). The goal of this study was to gain a closer insight into the optimal binding orientation of the two

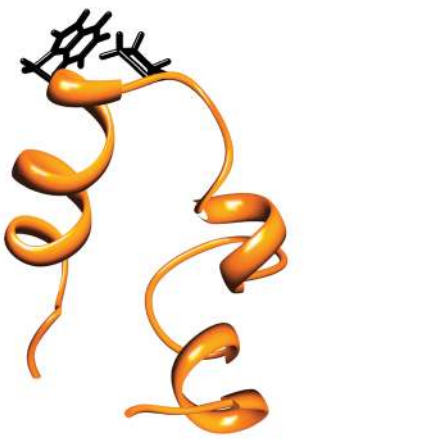


Fig. 36 NMR structure of native HP36 (PDB code 1VII)[217] showing the proximity between Pro62 and Trp64. The figure has been generated using UCSF Chimera.[57]

Fig. 37 NMR structure of the trp cage miniprotein (PDB code 1L2y, structure #0.1)[219] showing the side chain of Pro12. The figure has been generated using UCSF Chimera.[57]

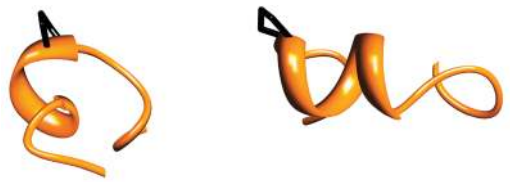


Fig. 38 NMR structure of IM1 (PDB code 1IM1, structure #0.1)[221] and crystal structure of Pn1A (PDB code 1PEN)[222] indicating the position of the substituted Pro. The figures were generated using UCSF Chimera.[57]

α-conotoxins by employing different substituents either of the *proS* or *proR* hydrogen at C4 of the prolyl side chain. However, the results for fluorinated analogues of both, Im1 and PnIA[A10L], were inconclusive in this regard. Fluorine, though slightly increasing helicity, has only little effect on the structure of the α-conotoxins and this result was independent of stereochemistry. Binding affinities and inhibitory activity of both α-conotoxins were slightly reduced upon fluorination but again, stereochemistry had no major impact.

These findings may be due to the fact that, in contrast to P4H that directly addresses the prolyl side chain, the receptor here does not have any direct interaction with the prolyl ring. Moreover, the minor structural perturbations imparted by fluorine in this context do not seem to be strong enough to evoke major effects. It may be interesting to note that although these residues are not conserved in α-conotoxins, Pn1A contains two additional prolines, one adjacent to Pro6 and the other within the C-terminal loop. It might be worth assessing effects of fluorination at these positions as well.

Single substitutions of proline residues can sometimes have pronounced effects on protein structure and activity. However, fluorine substitutions are often conservative regarding size and conformational biases and mostly seem to be tolerated without large structural effects. Whether an effect will be pronounced or almost negligible is hard to predict since many factors are involved. In turn, global substitutions of several or all Pro residues within a protein almost always have a measurable impact and, in many cases, they can be explained by invoking the intrinsic conformational properties of fluorinated proline analogues.

5.3 Global substitutions

5.3.1 Fibrillar proteins. Due to their interesting mechanical properties fibrillar proteins such as elastin and collagen are of special interest in biomedicine and biotechnology.[223,224] Both of these proteins contain repetitive sequences incorporating proline residues that are responsible for their unique folds as well as their mechanical properties. Fluorinated analogues of proline have been used to modify these proteins both to manipulate their macromolecular properties as well as to interrogate structural determinants of their folding.

Collagen—fluorinated Pro in polyproline II helices. It was known that the presence of Hyp within the ProHypGly repeat greatly stabilizes the collagen triple helix but the mechanism by which it exerts its stabilization remained mysterious for a long time. It was with the help of fluorinated analogues of Pro that the ''code for collagen's stability'' was finally deciphered by Raines and co-workers. Hyp was replaced by (4*R*)-FPro within the collagen model $(ProHypGly)_{10}$.[225,226] The overall stabilisation imparted by (4*R*)-FPro was shown to be twice as large as that of Hyp when compared to the parent $(ProProGly)_{10}$ triple helix. Based on the argument that carbon bound fluorine is a poor hydrogen bond acceptor that is unable to sustain a network of hydrogen bonds throughout the triple helix, the authors were able to disprove the hypothesis that collagen stability relies on bridging water molecules between neighbouring Hyp residues. In fact, the stabilisation they observed is based on fluorine's stereoelectronic effects on the proline ring pucker and *cis*/*trans* isomerization.[192] Although not undisputed (*vide infra*), the authors concluded that the same stabilization mechanism should apply to Hyp.

Most of our current knowledge about the stereoelectronic control of Pro conformation and *cis*/*trans* isomerization of Xaa–Pro peptide bonds by fluorine and other substituents was actually inspired by these studies. Following the pioneering work in the group of Raines several groups have investigated the role of stereoelectronic effects by substituting either the Xaa or Yaa position in the $(XaaYaaGly)_{7/10}$ repeat sequence with both diastereomers of fluoroproline. Crystal structures of collagen models[227] show that all peptide bonds in native collagen are *trans*. However, the ring pucker of Pro residues in the Xaa and Yaa position is different. While Xaa exhibits the C^{γ}-*endo* pucker, Yaa is found to be C^{γ}-*exo* puckered. It was hypothesized that by stereospecifically endowing Pro with (fluorine) substituents that preorganise the peptide bond and prolyl side chain, and thus make the backbone conform to the native structure, collagen analogues with greatly enhanced stability could be engineered. Compared to Pro and Hyp, (4*R*)-FPro more strongly favours the C^{γ}-*exo* pucker allowing intrastrand $n \rightarrow \pi^*$ interactions that stabilize *trans* peptide bonds. Thus, when incorporated into the Yaa position, (4*R*)-FPro strongly enhances collagen stability. In contrast, (4*S*)-FPro favours the C^{γ}-*endo* pucker and has a lower equilibrium constant for the *trans* to *cis* isomerization and therefore strongly destabilizes the triple helix.[192] The same result, albeit employing a different model, has been reported by the group of Luis Moroder.[228] Comparable studies have also been published for the Xaa position[229–231] that, as opposed to the Yaa position, adopts the C^{γ}-*endo* pucker. Because (4*S*)-FPro favours the same pucker it stabilizes the collagen triple helix at Xaa in contrast to (4*R*)-FPro which actually prevents triple helix formation when present at this position. Interestingly, a combination of stabilizing residues in models such as (4*S*)-FProHypGly-[231] or (4*S*)-FPro(4*R*)-FProGly-[195,232] repeats do not lead to hyperstable collagen analogues but rather strongly destabilize the trimer or even prevent its formation, respectively. Molecular modelling suggested that unfavourable steric OH···F/F···F interactions (Fig. 39) between adjacent strands in the triple helix may counterbalance or even override favourable stereoelectronic effects of fluorine.[195,233] The Raines group tested two alternative ways to yield stable triple helices with double fluorinated collagen mimics. Shifting the position of fluorine at the Xaa position to C3 in a way that also favours the C^{γ}-*endo* pucker, *i.e.* in $[(3S)\text{-FPro}(4R)\text{-FProGly}]_7$ they could show that a triple helix can be formed, by preventing steric clashes. However, the stability was substantially lower

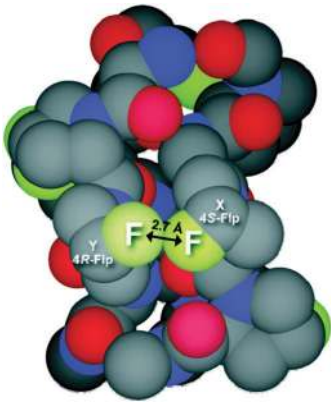


Fig. 39 Molecular model of the [(4*S*)-FPro(4*R*)-FProGly] repeat indicating the steric clash between fluorine atoms of adjacent strands. Note that the sum of van der Waals radii of two fluorine atoms is 2.94[234] (reproduced with permission from Hodges *et al.*[195] Copyright 2011, American Chemical Society).

than that of the (Pro(4*R*)-FProGly)$_7$ reference peptide. This was explained by C3-fluorine's inductive effect that weakens an interstrand $_{Pro}C{=}O\cdots HN_{Gly}$ hydrogen bond formed by the Xaa position. Because the Yaa position, on the other hand, is not involved in hydrogen bonds in collagen, this problem has not been observed here. Based on the argument that both 3-FPro diastereomers do not adopt suitable dihedral angles conforming to the requirements of the Yaa position, such analogues have never been investigated.

Hodges and Raines have also generated heterotrimeric assemblies of (ProProGly)$_7$ and [(4*S*)-FPro(4*R*)-FProGly]$_7$ showing that a 1 : 2 mixture indeed forms stable trimers, albeit with reduced stability compared to [(4*S*)-FProProGly]$_7$ and [Pro(4*R*)-FProGly]$_7$. Although they suggested a "code for strand association" as a model for the strand alignment in these heterotrimeric helices, it would be desirable to grow crystals to show how the individual strands associate.

It may be important to note that the applicability of fluoroprolines as probes to interrogate structural aspects of proline hydroxylation in collagen has been challenged by the fact that the mechanisms by which Hyp and FPro affect collagen stability when compared to Pro appear to be different.[235,236] However, the important point to make in the context of this review is that the effect of fluorination on collagen stability, for the most part, is an entropic one which highlights its impact on Pro conformation. To further prove that the entropic effect is not based on increasing hydrophobicity by fluorination, Shoulders *et al.* have also tested a host–guest collagen mimic containing a single Pro4^2-F$_2$ProGly sequence at a central position. The hypothesis was that if hydrophobicity is the stabilizing effect, it should stabilise the assembly more than (4*R*)-FPro.[210] However, 4^2-F$_2$Pro reduces stability compared to (4*R*)-FPro. Most importantly, it was shown to be as stable as native Pro at the same position that was readily explained by the fact that 4^2-F$_2$Pro and Pro exhibit the same conformational preferences. Furthermore, molecular modelling results showed that the double fluorinated variant does not suffer from steric strain caused by two vicinal fluorine atoms. A study in which Pro was globally substituted by 4^2-F$_2$Pro has not been pursued.

We note that in order to fully understand conformational aspects pertaining to the collagen structure and stability a more profound discussion that leads far beyond the scope of this review may be necessary. The reader is referred to a recent review by Shoulders and Raines[237] that provides an in-depth discussion of these aspects.

Elastin—fluorinated Pro in type II β-turns. Elastin consists of repetitive sequences of the pentad ValProGlyVal/IleGly that forms type II β-turns in which Pro occupies the (i + 1) position. Kim *et al.* have globally substituted all three C4-fluorinated Pro analogues for native Pro in the elastin-mimetic polypeptide $MGH_{10}S_2GHID_4KHM[(VPGVG)_4VPGIG]_{16}V$ using Pro-auxotrophic *E.coli* strains.[238] In their structural and thermodynamic investigations, however, the analogue containing 4,4-F$_2$Pro was not included, probably because its conformational properties are very similar to native Pro. Their actual goal was, to assess the role of Pro conformation during temperature induced, endothermic self-assembly of their elastin model peptide, by exerting the stereoelectronic control of single fluorine substitutions.[239] While (4*S*)-FPro (elastin 2) increased the midpoint of the transition by 8 K, it was decreased by 11.5 K by (4*R*)-FPro (elastin-3) compared to the native parent peptide elastin-1. CD-analyses and DSC measurements indicated that elastin-1 and -3 undergo two-state conformational transitions from a random-coil state to type II β-turns, while in the case of elastin-2 no such clear conclusion could be made. Importantly, the elastin-2 CD-spectra indicated the presence of alternative turn conformations. Moreover, elastin-3 containing (4*S*)-FPro was the only analogue that showed a measurable decrease in apparent heat capacity, indicating burial of hydrophobic surface area, *i.e.* aggregation of the peptide. NMR-investigations showed a detectable fraction of *cis*-Val-Pro peptide bonds in the case of elastin-2 which could be rationalized on the basis of the intrinsic conformational preferences of (4*S*)-FPro. However, the actual equilibrium population was less than 20%, which means that the macromolecular properties of the three elastin analogues should not exclusively correspond to the impact of fluorine on the proline ring pucker. Therefore, secondary effects on the main chain torsional angles ϕ and ψ were invoked to explain the observed differences. Previous crystal structures of cyclic elastin models had shown that Pro uniformly adopts a C$^\gamma$-*exo* ring pucker in type II β-turns in elastin.[240] However, it remained unclear whether this structural feature is persistent to the elastin repeat sequence or whether this was an artefact of the model or crystallization. The authors showed that by more strongly favouring the C$^\gamma$-*exo* pucker compared to native Pro, (4*R*)-FPro preorganizes the main chain to form the type II β-turn in elastin and thus decreases the transition temperature (and enthalpy) thereby proving that the C$^\gamma$-*exo* pucker is a required feature of elastin structure. However, the situation appeared more complex for the negative control (4*S*)-FPro. Based on CD-spectra the authors concluded that alternative turn structures may be present in elastin-2. Their theoretical studies (DFT calculations) applying both diastereomers of the model peptide Ac-FProGly-NHMe supported the conclusions made for (4*S*)-FPro but also showed that the intrinsic preference of (4*S*)-FPro for the C$^\gamma$-*endo* ring pucker is not reproduced in the

context of type II β-turns. However, the C^{γ}-*exo* isomer of (4*S*)-FPro is still less stable than the C^{γ}-*exo* isomer of (4*R*)-FPro in this context. Therefore, the energy of a type II β-turn is increased when directly comparing (4*S*)- to (4*R*)-FPro so that alternative turns, in particular the type I β-turn, become energetically more accessible.

5.3.2 Effects of fluorination on the PPII ⇄ PPI transition. Polyproline sequences such as $(Pro)_{10}$ adopt two different conformations (PPII and PPI) that strongly depend on solvent conditions. While in aqueous solution the left handed PPII helix in which all peptide bonds adopt the *trans* conformation is formed, the *cis* conformation is favoured in organic solvents such as *n*-propanol yielding the right handed PPI helix. It has been hypothesised that the *trans* conformation is stabilized by favourable $O_{i-1} \rightarrow C{=}O_i$ $n \rightarrow \pi^*$ interactions and thus the stereoelectronic effects of fluorine substituents could be used to control the polyproline structure.[241] Model peptides $(Pro)_{10}$GlyTyr were synthesised and Pro was globally replaced by (4*R*)- and (4*S*)-FPro as well as Hyp. Both (4*R*)-FPro and Hyp preferably adopt *trans* peptide bonds and were found to more strongly favour the PPII helix in both aqueous solvent and *n*-propanol, even at higher temperature. In turn, $[(4S)\text{-FPro}]_{10}$GlyTyr was shown to form a mixture of PPII and PPI in water, and, compared to Pro, its propensity to adopt a PPI helical conformation was enhanced in *n*-propanol. When comparing Pro, Hyp and its fluorinated analogues, solvation effects may have to be considered. Nevertheless, such effects should not play a key role in a direct comparison of both FPro diastereomers. Here, the findings may exclusively be explained by stereoelectronic effects of the fluorine atom. Because a (4*R*)-fluorine substituent preorganizes the C^{γ}-*exo* ring pucker that prefers the *trans* peptide bond, it enhances the proposed $n \rightarrow \pi^*$ interaction. When replacing the *proS* hydrogen such interactions are obviated by favouring a conformation that is more suitable for *cis* peptide bonds. Because Hyp and (4*R*)-FPro show similar behaviour, the mechanisms by which OH and F affect the polyproline structure can be explained by stereoelectronic control of Pro conformation.

To gain further insight, Chiang *et al.* pursued a kinetic study of the PPII ⇄ PPI transition using the host–guest model $(Pro)_5$-**Pro**-$(Pro)_5$GlyTyr instead of globally substituting all Pro residues.[242] Their kinetic data allowed conclusions about the effects of fluorine on the energetic barriers of both the PPI → PPII as well as the PPII → PPI folding transition. They found that electronegative substituents (not only fluorine but also OH and OMe) mostly accelerate the conversion form PPI (all-*cis*) to PPII (all-*trans*). The authors, therefore, concluded that stereoelectronic effects should not be primarily responsible for changing the energetic barrier of the PPI → PPII folding reaction. Moreover, the difference between (4*R*)-FPro and (4*S*)-FPro was not explained. Since the kinetic data do not account for uncertainties in their values, it is unclear whether these differences are relevant (the difference in $\Delta\Delta G^{\neq}$ between (4*R*) and (4*S*)-FPro was 0.3 kcal mol^{-1}). For the reverse reaction PPII → PPI they found that while (4*R*)-FPro increases the activation energy by roughly 0.7 kcal mol^{-1}, it is decreased by (4*S*)-FPro to the same extent when both are compared to Pro. The comparative interpretations in this paper are based on the assumption that all substitutions made did not affect the energy level of the transition state. The data covering the PPII → PPI conversion, however, support the finding that in, contrast to (4*S*)-FPro, $n \rightarrow \pi^*$ interactions that are favoured by (4*R*)-FPro stabilize *trans* prolyl peptide bonds and thus increase the energetic barrier to convert into a *cis* prolyl peptide bond. Based on these findings, the different effects of fluorinated prolines on polyproline structure seem to primarily stem from stereoelectronic effects on the energetic barrier of PPII → PPI conversion.

It has been hypothesized by Ruzza *et al.* that the stabilization of the PPII conformation could be useful to increase the binding affinity of protein ligands to SH3-domains, since the PPII conformation is a common feature of SH3-ligands.[243] They synthesized a series of peptides corresponding to the Pro-rich sequence of hematopoietic progenitor kinase 1 (HPK1) Pro**Pro**Pro**Leu**Pro**Pro**Lys**Pro**LysPhe in which they replaced Pro for either (4*R*)- or (4*S*)-FPro (the substitution positions in bold). They found that (4*S*)-FPro generally decreases the PPII helical content, while (4*R*)-FPro has the opposite effect. Applying CD-spectroscopy it was concluded that (4*R*)-FPro has a more pronounced effect when substituted for the $i + 1$ position of the $i \rightarrow (i + 3)$ PPI turn, *i.e.* positions 2, 5 and 8. Although explicit stability data were not reported, these findings are consistent when compared to $(Pro)_{10}$ sequences (*vide supra*). The hypothesis was that a stabilization of PPII conformation would decrease the entropic cost of binding to the SH3 domain thereby increasing the binding affinity. However, although the peptides containing (4*R*)-FPro are mostly better ligands, all other fluorinated analogues displayed decreased binding affinities compared to the parent peptide. In addition to the favourable entropic contribution, enthalpic changes induced by fluorine have to be considered. The authors noted that fluorine may weaken an important hydrogen bond between $_{Pro3}C{=}O$ and the $_{Tyr}OH$ of the target SH3 domain. Furthermore, aromatic–prolyl interactions seem to play a key role for binding. This latter aspect, however, has not been discussed in detail.

5.3.3 Globular proteins

Green fluorescent protein. Steiner *et al.* succeeded in expressing fluorinated variants of enhanced green fluorescent protein (EGFP) in which all 10 proline residues are replaced by 4-FPro diastereomers (Fig. 40).[244] While the protein containing (4*R*)-FPro was insoluble and unfolded, the (4*S*)-FPro variant displayed superior folding and fluorescent properties compared to the wild type EGFP. In the native protein nine out of ten Xaa-Pro peptide bonds are in the *trans* conformation and this structural feature is maintained in the fluorinated variant. However, despite its stronger preference for *trans* peptide bonds (4*R*)-FPro resulted in an unfolded protein. The finding that, instead, (4*S*)-FPro stabilized EGFP was initially surprising. One should remember that, although (4*S*)-FPro has a lower equilibrium constant for the *trans* → *cis* equilibrium it still prefers the *trans* over the *cis* conformation (see Section 5.1).

In addition, the *cis* conformation is favoured for the Met88-(4*S*)-FPro89 peptide bond, and since the isomerization is rate limiting, this single peptide bond can have a large impact.

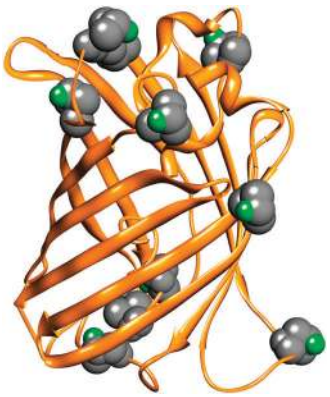


Fig. 40 Crystal structure of (4*S*)-FPro-EGFP (PDB code 2Q6P) showing the positions of (4*S*)-FPro.[244] The figure was generated using UCSF Chimera.[57]

In their crystal structure, the authors furthermore found that all but one (Pro56) adopt the predicted C^{γ}-*endo* pucker of (4*S*)-FPro. It is not clear, however, whether the ring pucker can be used to rationalize the stabilizing effect of (4*S*)-FPro since ring puckers could not be assigned in native EGFP. Nevertheless, the crystal structure revealed several close contacts between fluorine atoms and mostly backbone NH atoms. Some of these contacts do not seem to conform to ideal hydrogen bonding geometries and the little spatial separation of H and F may be structurally imposed rather than being a result of specific interactions. However, the fluorine atoms as well as the C^{γ}-*endo* pucker seem to be very well accommodated at these positions. As opposed to collagen or single poly-proline helices some of the proline residues are buried within the protein. Thus, it was argued that the increased hydrophobicity of the fluorinated proline residues may also be used to explain the stabilisation. However, five Pro side chains are solvent exposed and should thus have a contrary effect. It may also be worth mentioning that single fluorine substitutions of alkyl chains, though the C–F bond itself may not be very well solvated, do usually tend to decrease rather than increase hydrophobicity. A further in-depth thermodynamic study could help to elucidate the mechanism by which fluorination stabilizes EGFP. Moreover, single substitutions may be useful to study the role of the individual Pro residues and their conformational preferences in the context of EGFP.

KlenTaq DNA polymerase. Holzberger and Marx have studied the effect of globally replacing 32 Pro residues by 4-FPro on the stability and activity of *KlenTaq* DNA polymerase.[245] In contrast to EGFP, however, (4*S*)-FPro was not accepted by the expression system, although the same bacterial strain has been used for both proteins. Therefore, only a variant containing (4*R*)-FPro could be generated in which roughly 92% of all prolines were replaced. 21 of the replaced positions were identified and P701 was found to not have been substituted. In a case in which the level of unnatural amino acid incorporation is not 100% and not all the substituted positions can be identified, the product is very likely to be heterogeneous. Thus, all interpretations made in this paper very likely correspond to a mixture of proteins. Nevertheless, the authors found that, although (4*R*)-FPro tends to thermally destabilize the protein, its activity is practically identical to that of the wild type. Moreover, the error rate of the fluorinated polymerase remained unchanged. It is interesting, however, that, as opposed to the wild type, deletion errors were not detected when employing the fluorinated polymerase in a PCR based assay. This is probably one of the most interesting findings of this study.

6. Summary and overview

Table 3 provides an overview of all of the native peptides and proteins that have been reviewed here.

Fluorinated non-polar aliphatic and aromatic amino acids are the most frequently used analogues, which may in part be explained by their better synthetic accessibility compared to fluorinated polar and charged amino acids. Complete/multiple substitution of fluorinated aliphatic residues in the hydrophobic core of coiled coils always leads to an increase in stability. Because highly fluorinated amino acids bear significantly larger side chains, however, their global incorporation may affect the oligomerization state of such helical bundles. Such changes may seriously affect the function of antimicrobial peptides but these effects are rarely predictable and remain to be determined on a case-by-case basis. Global substitutions of fluorinated aliphatic residues in globular proteins, in contrast, often induce misfolding and loss of function. These findings may be explained by the fact that non-polar aliphatic residues are more evenly distributed throughout the structure of these proteins and are not only involved in formation of hydrophobic cores. Nevertheless, the combination of global fluorination and directed evolution holds promise as a method to generate fluorinated globular proteins with superior structural and functional properties. There is growing evidence that single substitutions within helical structures seem to destabilize folding, while the same substitutions within β-sheet rich structures lead to stabilization. In our opinion, however, it is still too early to draw general conclusions about secondary structural propensities of fluorinated amino acids, since many of them have never been investigated in this respect. Moreover, it is not easy to deconvolute steric and electronic effects even if high resolution structures are available. In this respect, computational chemistry has proven a highly valuable tool in interpreting experimental observations.

A great deal of literature considers proteins that contain fluorinated aromatic residues. Many of these studies, however, do not explicitly address the effects of fluorine. One of the most interesting applications of fluorinated aromatic amino acids is their use as probes to identify catalytically active residues, such as Tyr. Moreover, fluorination can be applied as a strategy to modulate the strength of cation–π interactions. In general, however, the effects of aromatic fluorine substitutions on protein structure and function are not easily rationalized. While single substitutions are mostly tolerated without major changes, global substitutions can have either favourable or unfavourable effects. Unfortunately, our current knowledge does not allow general conclusions to be made at this time.

Fluorinated analogues of charged and polar amino acids hold promise primarily as probes of catalytically active sites. By altering the chemical properties of these side chains, such as nucleophilicity and acidity/basicity, contributions of specific

Table 3 Overview of native peptides and proteins that have been modified with fluorinated analogues of the canonical amino acids

Amino acids	Fluorinated analogue	Modified peptide or protein	Major observations	Ref.
Hydrophobic, aliphatic				
Ala	3F-Ala	Cyclosporin A (for Val8)	80% native activity	19, 20
Val	4^3-F_3Val (diastereomeric)	NTL9(1–56) (for Val3 or 21 within the β-sheet)	Enhanced structural stability, increased folding rate and decreased unfolding rate	56
	4^3,$4'^3$-F_6Val	Gramicidin S (global for Val)	8–16 fold lower activity due to weaker membrane interaction	25
Ile	5^3-F_3Ile (diastereomeric)	mIL2	Marginal loss in proliferative activity	53
Leu	(4*S*)-5-FLeu	Ubiquitin (for core residues Leu43/67 or Leu50/67)	Native structure, marginal loss in stability	64
	4^2,$4'^2$-F_4Leu	GB1 (for internal β-strand position T53	Increased stability	59
	5^3-F_3Leu (diastereomeric)	Melittin (global for Leu)	Enhanced self-association within the membrane	24
		NTL9(1–56) (for Leu30 within the helix)	Destabilization (presumably due to unfavourable steric interactions)	58
		CAT (different levels of global substitution for Leu)	Enhanced secondary structure formation, stability decreases along with the extent of substitutions, misaggregation at elevated temperatures	65, 68
		GFP (global for Leu)	Loss of fluorescence	67
	5^3,$5'^3$-F_6Leu	Pexiganan (global for Leu and Ile)	Retained activity, haemolytic inactivity, improved membrane-bound proteolytic stability	23
		Buforin-2 (different positions)	Activity increases along with enhanced hydrophobicity, increased haemolytic activity	26
		Magainin-2 (different positions)	Similar/enhanced activity for partly fluorinated analogues, decreased activity for fully fluorinated variant due to enhanced self-assembly	26
		GLP-1 (different positions)	Substitutions around the cleavage site for DPP IV induce resistance towards degradation, substitutions at the binding site reduce affinity	51
		Protegrin-1 (positions 14 and 16 for Val)	Decreased antimicrobial activity (presumably due to lowered effective positive charge), different oligomerization, tetramers instead of dimers	61
		GB1 (for internal β-strand position 53)	Increased stability	59
Met	DfM	Alkaline protease A (for conserved Met214 and N-terminus)	Marginal decreased activity, N-terminal processing is blocked	77
	DfM, TfM	For-MetLeuPhe	Largely increased chemotactic activity	72
		Azurin (copper binding centre)	Reduction potential correlates with hydrophobicity of the copper binding site	78
		Bacteriophage lysozyme (LaL, global for Met)	Retained activity, marginal perturbation of rotational freedom of the SCH_3 group	73–75
	TfM	EGFP (two substitutions for Met)	N-terminal processing is blocked	76
		KlenTaq polymerase (global for Met)	Markedly decreased thermal stability	79
Aromatic				
Phe	2-FPhe, 3-FPhe, 4-FPhe	Human annexin V, *p. aeruginosa* azurin (global for Phe)	Fluorinated protein exhibits two characteristic fingers in UV spectra (260–270 nm)	84
		Restriction endonuclease *Pvu*II (global for Phe)	3F-Phe 2 fold higher activity than WT, 2F-and 4F mutants lower activity than WT	85
		Histone acetyltransferase tGN5 (global for Phe)	Loss in secondary structure and reduced enzymatic resistance	86
	4-FPhe	*Candida antractica* lipase B (CalB) (global for Phe)	Prolonged shelf life	87
	2,3,4,5,6-F_5Phe	Aβ(16–22) [Ac-KLVFFAE-NH_2] (single and double substitution of Phe)	Enhanced self-assembly with additive effects for double mutant	89
		Villin headpiece (positions 6, 10, and 17 for Phe)	One of seven mutants more stably folded than WT	81, 95
	2,3,4,5-F_5Phe; 2,3,5,6-F_5Phe	Villin headpiece (positions 6, 10, and 17 for Phe)	2,3,4,5-F_5Phe and 2,3,5,6-F_5Phe mutants more stable than 2,3,4,5,6-F_5Phe mutant	100
	4-FPhe, 3,5-F_2Phe, 3,4,5-F_3Phe	Glycine receptor (positions 63, 99, 159, and 207 for Phe)	Identification of a cation–π-interaction with position 159 in receptor-agonist binding	107, 108
		Sodium channel $Na_V1.4$ (position 401 for Phe)	Identification of a cation–π-interaction in receptor–toxin binding	110
	3-FPhe, 4-FPhe, 3,4-F_2Phe, 2,3,4,5,6-F_5Phe	Tridecapeptide α-factor (position 13 for Tyr)	Identification of a cation–π-interaction in receptor–agonist binding	111
	2,3-F_2Phe	Nonapeptide from Wilms tumor protein 1 (position 1 for Ser)	Mutant induced T cell response similar to WT	112
	3-FPhe, 4-FPhe, 3,4-F_2Phe	Human T-cell lymphotropic virus Tax peptide (position 5 for Tyr)	Increased receptor binding with additive effects for 3,4-F_2Phe mutant	114

Table 3 (*continued*)

Amino acids	Fluorinated analogue	Modified peptide or protein	Major observations	Ref.
Tyr	2-FTyr, 3-FTyr, 2,3,5,6-F_4Tyr	Nicotinic acetylcholine receptor (positions 93, 190, and 193 for Tyr)	Analogues with 2-FTyr or 3-FTyr at position 93 slightly more potent than WT	125
	2-FTyr, 3-FTyr, 2,3,5,6-F_4Tyr	Staphylococcal nuclease (positions 27 and 93 for Tyr)	Correlation between stability and pK_a of tyrosine's OH-group	126, 127
		Glutathione transferase (position 9 for Tyr)	Loss of catalytic efficiency	128
	2,3-F_2Tyr, 3,5-F_2Tyr, 2,3,5,6-F_4Tyr	*E.coli* ribonucleotide reductase (position 122 for Tyr)	Identification of Tyr as a redox active amino acids in protein	118, 132
	2-FTyr, 3-FTyr, 2,3-F_2Tyr	Δ^5-3-ketosteroid isomerase (position 14 for Tyr)	3F-Tyr and 2,3F_2-Tyr effectuate loss in catalytic efficiency; 2F-Tyr variant is more efficient	134
	3-FTyr	Human annexin V (global substitution)	Reduced thermal stability	124
		P. aeruginosa azurin (global substitution)	Increased thermal stability	
		Organophosphate hydrolase (global substitution)	Fluorinated mutants have extended pH-optimum of action	136
		Human manganese superoxide dismutase (global substitution)	Reduced catalytic activity	138
		Vaccinia virus DNA topoisomerase I (position 274 for Tyr)	Reduced enzymatic activity	140
		Human topoisomerase I (position 723 for Tyr)	Reduced enzymatic activity	141
		Photosystem II (global substitution)	Reduced activation energy	144
	2-FTyr, 3-FTyr	EGFP (global substitution)	Fluorescence emission of 3-FTyr mutant is red shifted and its intensity is decreased by 10%; fluorescence emission of 2-FTyr analogue is blue shifted	151, 152
	3-FTyr, 2,6-F_2Tyr	GFP(position 66 for Tyr)	Fluorescence emission of 3F-Tyr mutant is red shifted; fluorescence emission of 2,6-F_2Tyr analogue is blue shifted	154
	3-FTyr	DsRed (global substitution)	Fluorescence emission of 3-FTyr mutant is blue shifted	156
Trp	4-FTrp, 5-FTrp, 6-FTrp	Staphylococcal nuclease (position 140 for Trp)	5-FTrp and 6-FTrp have no effect on structural stability; 4-FTrp has a modest stabilizing effect; phosphorescence spectra of mutant is red shifted	158, 164
		Annexin V (position 187 for Trp)	5-FTrp increases thermal stability, 4F-Trp and 6-FTrp induce lower thermal stability	166
	4-FTrp	*E.coli* arginyl-TRNA synthase (global substitution)	First order rate constant is decreased; no significant effect on thermal stability, slight blue shift in UV spectra	168
	4-FTrp, 5-FTrp, 6-FTrp, 7-FTrp	ECFP (global substitution)	4-FTrp, 5-FTrp, and 7-FTrp blue shift in emission and absorption maxima	169
	4-FTrp , 5-FTrp, 6-FTrp, 5,7-F_2Trp, 5,6,7-F_3Trp, 4,5,6,7-F_4Trp	Nicotinic acetylcholine receptor (positions, α86, α184, α149 and γ55 for Trp)	Identification of a cation–π-interaction with Trpα149 in receptor–agonist binding	106
	4-FTrp, 5-FTrp, 6-FTrp	Streptavidin (SAV) (global substitution)	5-FTrp-SAV predicted to undergo more stable hydrogen bonding network when binding to biotin than WT	170
Charged, Polar				
His	(2F)-His	Antrax toxin (protective antigen)	Pore formation is not hampered (indicates that protonation of His does not trigger pore formation)	178, 179
	(4F)-His	Ribonuclease A (for either or both of two His at the catalytic site, His12 and His119)	Small loss of catalytic efficiency (proof that proton transfer is not rate limiting to the cleavage reaction)	177
	(2F)-His, (4F)-His)	chaperone PapD	No structural or functional studies reported	180
Asn	3-FAsn (diastereomeric)	**Asn**LeuThr glycosylation motif	Enzymatic glycosylation is blocked	181–183
Proline	3-FPro (diastereomeric)	Z-GlyPhe**Pro**Gly-OEt	Enzymatic hydroxylation at C4 is not affected	201
	(3*S*)-FPro, (4*R*)-FPro	Collagen model (**Xaa****Yaa**Gly)$_7$ → global substitution for Xaa by (3*S*)-FPro and Yaa by (4*R*)-FPro)	Reduced stability compared to the (Pro(4*R*)FProGly)$_7$ reference	195
	(4*R*)-FPro, (4*S*)-FPro	PEG-GlyTyr**Pro**Gly-OEt	No enzymatic hydroxylation of (4*R*)-FPro, (4*S*)-FPro is converted at reduced rate to give 4-ketoproline as the final product	202
		Type IV collagen host guest model (single substitution)	Compared to Hyp, (4*R*)-FPro stabilizes and (4*S*)-FPro destabilizes the triple helix; activity to promote cell spreading and signalling correlates with activity	208
		N-terminal domain of minicollagen-1 (single substitution of Pro24)	Enhanced accumulation of a folding intermediate with (4*R*)-FPro, only traces of the intermediate for (4*S*)-FPro	211

Table 3 (*continued*)

Amino acids	Fluorinated analogue	Modified peptide or protein	Major observations	Ref.
	(4*R*)-FPro, (4*S*)-FPro	Proctolin (ArgTyrLeu**Pro**Thr)	Intrinsic preference of (4*R*)-FPro and (4*S*)-FPro to adopt either *cis* or *trans* peptide bonds is not reflected (reversal);	212
		Villin headpiece subdomain (single substitution of Pro62)	(4*R*)-FPro destabilizes while (4*S*)-FPro stabilizes folding	216
		Trp cage miniprotein	Stabilization by (4*R*)-FPro, destabilization by (4*S*)-FPro	218
		α-conotoxinIM1, α-conotoxin PnIA[A10L] (for conserved Pro6)	Slightly increased helicity, slightly reduced activity	220
		Elastin (global substitution)	Stabilization by (4*S*)-FPro, destabilization by (4*R*)-FPro	238
		$(Pro)_{10}$-GlyTyr (global substitution)	(4*R*)-FPro favours PPII helix formation, (4*S*)-FPro favours PPI helix formation	241
		Hematopoietic progenitor kinase 1 (multiple substitutions)	(4*R*)-FPro favours PPII helix formation, (4*S*)-FPro favours PPI helix formation, SH3 binding affinity: native > (4*R*)-FPro > (4*S*)-FPro	243
		EGFP (global substitution)	(4*R*)-FPro: insoluble aggregates, (4*S*)-FPro: superior folding and fluorescence	244
		KlenTaq Polymerase (global substitution)	(4R)-FPro: thermal destabilization, retained activity	245
	(4*R*)-FPro	AcGly-(Pro**Hyp**Gly)$_8$-Gly-NH_2 (single substitution within the fourth repeat	Destabilization of the triple helix compared to Pro and Hyp	209
		Ac-(Pro**Pro**Gly)$_{10}$-NH_2 (single substitution within the fifth repeat)	Stabilization compared to Pro	210
	(4*R*)-FPro, (4*S*)-FPro, 4^2-F_2Pro	Suc-AlaSer**Pro**Phe-p-nitroanilide barstar(Cys40Ala/Cys82Ala/Pro27Ala)	Catalytic efficiency of peptidyl *cis*/*trans* isomerases depends on stereochemistry of 4-FPro, enzymes are inefficient to accelerate isomerization of peptide bonds containing 4^2-F_2Pro	206
		Collagen models (**Xaa****Yaa**Gly)$_{7/10}$ → global substitution for Yaa	Stabilization by (4*R*)-FPro, destabilization by (4*S*)-FPro	192, 225, 226, 229–231
		→ Global substitution for Xaa by (4*S*)-FPro and Yaa by (4*R*)-FPro)	Large destabilization due to unfavourable steric interactions of the fluorine atoms	195, 232
	4^2-F_2Pro	→ Single substitution of Yaa	Equally stable as the analogue containing Pro	210

residues like His can be investigated. However, studies dealing with fluorinated charged and polar residues are rare, presumably due to their complex synthesis.

Methionine and proline have a lower natural abundance and are, therefore, less frequently used. Fluorinated Met analogues can be used as NMR probes, but only a few studies have determined the structural and functional consequences of Met fluorination. Therefore, general conclusions cannot be drawn. The results gained with fluorinated Pro analogues can often be interpreted on the basis of the intrinsic conformational properties of the Pro analogue under consideration. For example, several studies employed fluorinated Pro analogues as structural probes to assess the conformational biases of protein structure. As such, fluorinated Pro has been used to unravel the origin of collagen's structure and stability. Moreover, these analogues have been used to prove the structural importance of *cis*/*trans* isomerisation and $^{Pro}CH\cdots\pi$ interactions. By choosing an analogue with appropriate conformational properties, the structural stability of the peptide or protein under consideration can often be enhanced. However, judging from the results published to date, it seems that interactions with enzymes involving the Pro side chain cannot usually be improved.

7. Conclusions

Protein engineering both by biological and synthetic approaches serves two important goals: (1) understanding protein structure formation and function and (2) improving these properties. These goals can be achieved either by using the canonical amino acids or by incorporating non-natural analogues. As shown in this review, fluorinated amino acids have become very prominent tools in the protein engineering field.

Our major objective in summarizing the results obtained in this field since the mid-90s was to generalize the effects of introducing fluorinated amino acids into naturally occurring protein sequences. With respect to their impact on structure, it is possible to state that these building blocks are largely non-perturbing. However, in spite of the fact that their intrinsic physicochemical properties are well understood, the way in which they influence protein stability (thermal and enzymatic) and function are not easily generalized or rationalized. In some cases, protein properties can be improved, as has been shown in fluorinated systems with greater thermal or proteolytic stability. On the other hand, fluorinated amino acids often have deleterious effects on the biological function of the enzyme being studied. For that reason, we conclude that further studies in this field are essential for contributing to our knowledge base to enable the rational design of non-natural proteins, so that our successful discoveries are not left to chance.

Acknowledgements

The authors are grateful to the DFG (Graduate School #1582/1 "Fluorine as a key element") for financial support, to Dr. Allison Ann Berger for fruitful discussions and Dr. Pamela Winchester for proof reading of the manuscript.

References

1 S. Purser, P. R. Moore, S. Swallow and V. Gouverneur, *Chem. Soc. Rev.*, 2008, **37**, 320.
2 C. P. R. Hackenberger and D. Schwarzer, *Angew. Chem., Int. Ed.*, 2008, **47**, 10030.
3 T. S. Young and P. G. Schultz, *J. Biol. Chem.*, 2010, **285**, 11039.
4 U. Arnold, *Biotechnol. Lett.*, 2009, **31**, 1129.
5 A. Sutherland and C. L. Willis, *Nat. Prod. Rep.*, 2000, **17**, 621.
6 X.-L. Qiu, W.-D. Meng and F.-L. Qing, *Tetrahedron*, 2004, **60**, 6711.
7 R. Smits, C. D. Cadicamo, K. Burger and B. Koksch, *Chem. Soc. Rev.*, 2008, **37**, 1727.
8 R. Smits and B. Koksch, *Curr. Top. Med. Chem.*, 2006, **6**, 1483.
9 A. S. Ulrich, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2005, **46**, 1.
10 S. L. Cobb and C. D. Murphy, *J. Fluorine Chem.*, 2009, **130**, 132.
11 D. O'Hagan, *Chem. Soc. Rev.*, 2008, **37**, 308.
12 R. Berger, G. Resnati, P. Metrangolo, E. Weber and J. Hulliger, *Chem. Soc. Rev.*, 2011, **40**, 3496.
13 S. A. Samsonov, M. Salwiczek, G. Anders, B. Koksch and M. T. Pisabarro, *J. Phys. Chem. B*, 2009, **113**, 16400.
14 G. Bott, L. D. Field and S. Sternhell, *J. Am. Chem. Soc.*, 1980, **102**, 5618.
15 T. Nagai, G. Nishioka, M. Koyama, A. Ando, T. Miki and I. Kumadaki, *J. Fluorine Chem.*, 1992, **57**, 229.
16 H. M. Walborsky and J. H. Lang, *J. Am. Chem. Soc.*, 1956, **78**, 4314.
17 M. Schlosser, *Angew. Chem., Int. Ed.*, 1998, **37**, 1496.
18 A. K. Mitra, I. Ostashevsky and C. F. Brewer, *Int. J. Pept. Protein Res.*, 1983, **22**, 495.
19 O. D. Hensens, R. F. White, R. T. Goegelman, E. S. Inamine and A. A. Patchett, *J. Antibiot.*, 1992, **45**, 133.
20 A. A. Patchett, D. Taub, O. D. Hensens, R. T. Goegelman, L. Yang, F. Dumont, L. Peterson and N. H. Sigal, *J. Antibiot.*, 1992, **45**, 94.
21 E. Dessipri and D. A. Tirrell, *Macromolecules*, 1994, **27**, 5463.
22 H. R. Hoveyda and J.-F. Pinault, *Org. Lett.*, 2006, **8**, 5849.
23 L. M. Gottler, H. Y. Lee, C. E. Shelburne, A. Ramamoorthy and E. N. G. Marsh, *ChemBioChem*, 2008, **9**, 370.
24 A. Niemz and D. A. Tirrell, *J. Am. Chem. Soc.*, 2001, **123**, 7407.
25 T. Arai, T. Imachi, T. Kato, H. I. Ogawa, T. Fujimoto and N. Nishino, *Bull. Chem. Soc. Jpn.*, 1996, **69**, 1383.
26 H. Meng and K. Kumar, *J. Am. Chem. Soc.*, 2007, **129**, 15615.
27 D. N. Woolfsona. J. M. S. David and A. D. Parry, in *Advances in Protein Chemistry*, Academic Press, 2005, vol. 70, p. 79.
28 Y. Tang, G. Ghirlanda, N. Vaidehi, J. Kua, D. T. Mainz, W. A. Goddard, W. F. DeGrado and D. A. Tirrell, *Biochemistry*, 2001, **40**, 2790.
29 S. Son, I. C. Tanrikulu and D. A. Tirrell, *ChemBioChem*, 2006, **7**, 1251.
30 P. Wang, A. Fichera, K. Kumar and D. A. Tirrell, *Angew. Chem., Int. Ed.*, 2004, **43**, 3664.
31 Y. Tang, G. Ghirlanda, W. A. Petka, T. Nakajima, W. F. DeGrado and D. A. Tirrell, *Angew. Chem., Int. Ed.*, 2001, **40**, 1494.
32 Y. Tang and D. A. Tirrell, *J. Am. Chem. Soc.*, 2001, **123**, 11089.
33 J. K. Montclare, S. Son, G. A. Clark, K. Kumar and D. A. Tirrell, *ChemBioChem*, 2009, **10**, 84.
34 B. Bilgiçer, A. Fichera and K. Kumar, *J. Am. Chem. Soc.*, 2001, **123**, 4393.
35 B. Bilgiçer, X. Xing and K. Kumar, *J. Am. Chem. Soc.*, 2001, **123**, 11815.
36 B. Bilgiçer and K. Kumar, *Tetrahedron*, 2002, **58**, 4105.
37 B. Bilgiçer and K. Kumar, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 15324.
38 N. Naarmann, B. Bilgiçer, H. Meng, K. Kumar and C. Steinem, *Angew. Chem., Int. Ed.*, 2006, **45**, 2588.
39 K.-H. Lee, H.-Y. Lee, M. M. Slutsky, J. T. Anderson and E. N. G. Marsh, *Biochemistry*, 2004, **43**, 16277.
40 H.-Y. Lee, K.-H. Lee, H. M. Al-Hashimi and E. N. G. Marsh, *J. Am. Chem. Soc.*, 2006, **128**, 337.
41 L. M. Gottler, R. de la Salud-Bea and E. N. G. Marsh, *Biochemistry*, 2008, **47**, 4484.
42 P. Luo and R. L. Baldwin, *Biochemistry*, 1997, **36**, 8413.
43 D. Roccatano, G. Colombo, M. Fioroni and A. E. Mark, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 12179.
44 B. C. Buer, R. de la Salud-Bea, H. M. Al Hashimi and E. N. G. Marsh, *Biochemistry*, 2009, **48**, 10810.
45 S. S. Pendley, Y. B. Yu and T. E. Cheatham, *Proteins: Struct. Funct. Bioinf.*, 2009, **74**, 612.
46 C. Jäckel, W. Seufert, S. Thust and B. Koksch, *ChemBioChem*, 2004, **5**, 717.
47 C. Jäckel, M. Salwiczek and B. Koksch, *Angew. Chem., Int. Ed.*, 2006, **45**, 4198.
48 M. Salwiczek, S. Samsonov, T. Vagt, E. Nyakatura, E. Fleige, J. Numata, H. Colfen, M. T. Pisabarro and B. Koksch, *Chem.–Eur. J.*, 2009, **15**, 7628.
49 M. Salwiczek, *Dissertation*, Freie Universität, Berlin, 2010.
50 M. Salwiczek and B. Koksch, *ChemBioChem*, 2009, **10**, 2867.
51 H. Meng, S. T. Krishnaji, M. Beinborn and K. Kumar, *J. Med. Chem.*, 2008, **51**, 7303.
52 R. Paulini, K. Müller and F. Diederich, *Angew. Chem., Int. Ed.*, 2005, **44**, 1788.
53 P. Wang, Y. Tang and D. A. Tirrell, *J. Am. Chem. Soc.*, 2003, **125**, 6900.
54 H.-P. Chiu, Y. Suzuki, D. Gullickson, R. Ahmad, B. Kokona, R. Fairman and R. P. Cheng, *J. Am. Chem. Soc.*, 2006, **128**, 15556.
55 H.-P. Chiu and R. P. Cheng, *Org. Lett.*, 2007, **9**, 5517.
56 J.-C. Horng and D. P. Raleigh, *J. Am. Chem. Soc.*, 2003, **125**, 9286.
57 E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605.
58 B. Anil, S. Sato, J.-H. Cho and D. P. Raleigh, *J. Mol. Biol.*, 2005, **354**, 693.
59 H.-P. Chiu, B. Kokona, R. Fairman and R. P. Cheng, *J. Am. Chem. Soc.*, 2009, **131**, 13192.
60 T. Gallagher, P. Alexander, P. Bryan and G. L. Gilliland, *Biochemistry*, 1994, **33**, 4721.
61 L. M. Gottler, R. de la Salud Bea, C. E. Shelburne, A. Ramamoorthy and E. N. G. Marsh, *Biochemistry*, 2008, **47**, 9243.
62 R. Mani, M. Tang, X. Wu, J. J. Buffy, A. J. Waring, M. A. Sherman and M. Hong, *Biochemistry*, 2006, **45**, 8341.
63 L. M. Gottler, R. de la Salud Bea, C. E. Shelburne, A. Ramamoorthy and E. N. G. Marsh, *Biochemistry*, 2008, **47**, 9243.
64 D. Alexeev, P. N. Barlow, S. M. Bury, J. D. Charrier, A. Cooper, D. Hadfield, C. Jamieson, S. M. Kelly, R. Layfield, R. J. Mayer, H. McSparron, N. C. Price, R. Ramage, L. Sawyer, B. A. Starkmann, D. Uhrin, J. Wilken and D. W. Young, *ChemBioChem*, 2003, **4**, 894.
65 T. Panchenko, W. W. Zhu and J. K. Montclare, *Biotechnol. Bioeng.*, 2006, **94**, 921.
66 E. Neil and G. Marsh, *Chem. Biol.*, 2000, **7**, R153.
67 T. H. Yoo, A. J. Link and D. A. Tirrell, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 13887.
68 J. K. Montclare and D. A. Tirrell, *Angew. Chem., Int. Ed.*, 2006, **45**, 4518.
69 C. Jäckel, P. Kast and D. Hilvert, *Annu. Rev. Biophys.*, 2008, **37**, 153.
70 D. O'Hagan and H. S. Rzepa, *Chem. Commun.*, 1997, 645.
71 H. S. Duewel and J. F. Honek, *J. Protein Chem.*, 1998, **17**, 337.
72 M. E. Houston Jr., L. Harvath and J. F. Honek, *Bioorg. Med. Chem. Lett.*, 1997, **7**, 3007.
73 H. Duewel, E. Daub, V. Robinson and J. F. Honek, *Biochemistry*, 1997, **36**, 3404.
74 H. S. Duewel, E. Daub, V. Robinson and J. F. Honek, *Biochemistry*, 2001, **40**, 13167.
75 P. Cleve, V. Robinson, H. S. Duewel and J. F. Honek, *J. Am. Chem. Soc.*, 1999, **121**, 8475.
76 N. Budisa, O. Pipitone, I. Siwanowicz, M. Rubini, P. P. Pal, T. A. Holak and M. L. Gelmi, *Chem. Biodiversity*, 2004, **1**, 1465.
77 P. Walasek and J. Honek, *BMC Biochem.*, 2005, **6**, 21.
78 D. K. Garner, M. D. Vaughan, H. J. Hwang, M. G. Savelieff, S. M. Berry, J. F. Honek and Y. Lu, *J. Am. Chem. Soc.*, 2006, **128**, 15608.
79 B. Holzberger, M. Rubini, H. M. Möller and A. Marx, *Angew. Chem., Int. Ed.*, 2010, **49**, 1324.
80 M. A. Danielson and J. J. Falke, *Annu. Rev. Biophys. Biomol. Struct.*, 1996, **25**, 163.

81 M. G. Woll, E. B. Hadley, S. Mecozzi and S. H. Gellman, *J. Am. Chem. Soc.*, 2006, **128**, 15932.
82 T. Tsushima, K. Kawada, J. Nishikawa, T. Sato, K. Tori, T. Tsuji and S. Misaki, *J. Org. Chem.*, 1984, **49**, 1163.
83 M. Schlosser, N. Brügger, W. Schmidt and N. Amrhein, *Tetrahedron*, 2004, **60**, 7731.
84 C. Minks, R. Huber, L. Moroder and N. Budisa, *Anal. Biochem.*, 2000, **284**, 29.
85 M. A. Dominguez, K. C. Thornton, M. G. Melendez and C. M. Dupureur, *Proteins: Struct. Funct. Bioinf.*, 2001, **45**, 55.
86 N. Voloshchuk, A. Y. Zhu, D. Snydacker and J. K. Montclare, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 5449.
87 N. Budisa, W. Wenger and B. Wiltschi, *Mol. BioSyst.*, 2010, **6**, 1630.
88 S. M. Butterfield, P. R. Patel and M. L. Waters, *J. Am. Chem. Soc.*, 2002, **124**, 9751.
89 F. T. Senguen, N. R. Lee, X. Gu, D. M. Ryan, T. M. Doran, E. A. Anderson and B. L. Nilsson, *Mol. BioSyst.*, 2010, **7**, 486.
90 P. P. Geurink, N. Liu, M. P. Spaans, S. L. Downey, A. M. C. H. van den Nieuwendijk, G. A. van der Marel, A. F. Kisselev, B. I. Florea and H. S. Overkleeft, *J. Med. Chem.*, 2010, **53**, 2319.
91 L. Borissenko and M. Groll, *Chem. Rev.*, 2007, **107**, 687.
92 M. L. Waters, *Curr. Opin. Chem. Biol.*, 2002, **6**, 736.
93 M. L. Waters, *Biopolymers*, 2004, **76**, 435.
94 C. A. Hunter and J. K. M. Sanders, *J. Am. Chem. Soc.*, 1990, **112**, 5525.
95 G. Cornilescu, E. B. Hadley, M. G. Woll, J. L. Markley, S. H. Gellman and C. C. Cornilescu, *Protein Sci.*, 2007, **16**, 14.
96 H. Zheng, K. Comeforo and J. Gao, *J. Am. Chem. Soc.*, 2008, **131**, 18.
97 T. K. Chiu, J. Kubelka, R. Herbst-Irmer, W. A. Eaton, J. Hofrichter and D. R. Davies, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 7517.
98 F. Ponzini, R. Zagha, K. Hardcastle and J. S. Siegel, *Angew. Chem., Int. Ed.*, 2000, **39**, 2323.
99 A. P. West, S. Mecozzi and D. A. Dougherty, *J. Phys. Org. Chem.*, 1997, **10**, 347.
100 H. Zheng and J. Gao, *Angew. Chem., Int. Ed.*, 2010, **49**, 8635.
101 H. Robson Marsden, J. G. E. M. Fraaije and A. Kros, *Angew. Chem., Int. Ed.*, 2010, **49**, 8570.
102 M. A. Cejas, W. A. Kinney, C. Chen, J. G. Vinter, H. R. Almond, K. M. Balss, C. A. Maryanoff, U. Schmidt, M. Breslav, A. Mahan, E. Lacy and B. E. Maryanoff, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 8513.
103 B. Nieswandt and S. P. Watson, *Blood*, 2003, **102**, 449.
104 N. Zacharias and D. A. Dougherty, *Trends Pharmacol. Sci.*, 2002, **23**, 281.
105 J. P. Gallivan and D. A. Dougherty, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 9459.
106 W. Zhong, J. P. Gallivan, Y. Zhang, L. Li, H. A. Lester and D. A. Dougherty, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 12088.
107 S. A. Pless, K. S. Millen, A. P. Hanek, J. W. Lynch, H. A. Lester, S. C. R. Lummis and D. A. Dougherty, *J. Neurosci.*, 2008, **28**, 10937.
108 S. A. Pless, A. P. Hanek, K. L. Price, J. W. Lynch, H. A. Lester, D. A. Dougherty and S. C. R. Lummis, *Mol. Pharmacol.*, 2011, **79**, 742.
109 L. Q. Chen, M. Chahine, R. G. Kallen, R. L. Barchi and R. Horn, *FEBS Lett.*, 1992, **309**, 253.
110 V. P. Santarelli, A. L. Eastwood, D. A. Dougherty, R. Horn and C. A. Ahern, *J. Biol. Chem.*, 2007, **282**, 8044.
111 S. Tantry, F.-X. Ding, M. Dumont, J. M. Becker and F. Naider, *Biochemistry*, 2010, **49**, 5007.
112 M. Gómez-Nuñez, K. J. Haro, T. Dao, D. Chau, A. Won, S. Escobar-Alvarez, V. Zakhaleva, T. Korontsvit, D. Y. Gin and D. A. Scheinberg, *PLoS One*, 2008, **3**, e3938.
113 E. S. Doubrovina, M. M. Doubrovin, S. Lee, J. H. Shieh, G. Heller, E. Pamer and R. J. O'Reilly, *Clin. Cancer Res.*, 2004, **10**, 7207.
114 K. H. Piepenbrink, O. Y. Borbulevych, R. F. Sommese, J. Clemens, K. M. Armstrong, C. Desmond, P. Do and B. M. Baker, *Biochem. J.*, 2009, **423**, 353.
115 Y. H. Ding, K. J. Smith, D. N. Garboczi, U. Utz, W. E. Biddison and D. C. Wiley, *Immunity*, 1998, **8**, 403.
116 P. A. Cole, D. Sondhi and K. Kim, *Pharmacol. Ther.*, 1999, **82**, 219.
117 K. Kim and P. A. Cole, *J. Am. Chem. Soc.*, 1998, **120**, 6851.
118 M. R. Seyedsayamdost, S. Y. Reece, D. G. Nocera and J. Stubbe, *J. Am. Chem. Soc.*, 2006, **128**, 1569.
119 C. Niemann and M. M. Rapport, *J. Am. Chem. Soc.*, 1946, **68**, 1671.
120 T. J. McCord, D. R. Smith, D. W. Winters, J. F. Grimes, K. L. Hulme, L. Q. Robinson, L. D. Gage and A. L. Davis, *J. Med. Chem.*, 1975, **18**, 26.
121 A. Weissman and B. K. Koe, *J. Pharmacol. Exp. Ther.*, 1967, **155**, 135.
122 R. M. Hochster and J. H. Quastel, *Metabolic Inhibitors*, Academic Press, New York, 1963.
123 I. Baumel, G. P. Carlson and H. Lal, *Brain Res.*, 1971, **28**, 357.
124 C. Minks, S. Alefelder, L. Moroder, R. Huber and N. Budisa, *Tetrahedron*, 2000, **56**, 9431.
125 P. C. Kearney, M. W. Nowak, W. Zhong, S. K. Silverman, H. A. Lester and D. A. Dougherty, *Mol. Pharmacol.*, 1996, **50**, 1401.
126 J. S. Thorson, E. Chapman, E. C. Murphy, P. G. Schultz and J. K. Judice, *J. Am. Chem. Soc.*, 1995, **117**, 1157.
127 J. S. Thorson, I. Shin, E. Chapman, G. Stenberg, B. Mannervik and P. G. Schultz, *J. Am. Chem. Soc.*, 1998, **120**, 451.
128 Y.-J. Zheng and R. L. Ornstein, *J. Am. Chem. Soc.*, 1997, **119**, 1523.
129 J. F. Parsons and R. N. Armstrong, *J. Am. Chem. Soc.*, 1996, **118**, 2295.
130 G. Xiao, J. F. Parsons, R. N. Armstrong and G. L. Gilliland, *J. Am. Chem. Soc.*, 1997, **119**, 9325.
131 G. Xiao, J. F. Parsons, K. Tesh, R. N. Armstrong and G. L. Gilliland, *J. Mol. Biol.*, 1998, **281**, 323.
132 M. R. Seyedsayamdost, C. S. Yee, S. Y. Reece, D. G. Nocera and J. Stubbe, *J. Am. Chem. Soc.*, 2006, **128**, 1562.
133 C. S. Yee, M. C. Y. Chang, J. Ge, D. G. Nocera and J. Stubbe, *J. Am. Chem. Soc.*, 2003, **125**, 10506.
134 B. Brooks, R. S. Phillips and W. F. Benisek, *Biochemistry*, 1998, **37**, 9738.
135 Y. Park, J. Luo, P. G. Schultz and J. F. Kirsch, *Biochemistry*, 1997, **36**, 10517.
136 G. J. Klarmann, B. M. Eisenhauer, Y. Zhang, K. Sitaraman, D. K. Chatterjee, S. M. Hecht and S. F. J. Le Grice, *Protein Expression Purif.*, 2004, **38**, 37.
137 Y. A. Votchitseva, E. N. Efremenko and S. Varfolomeyev, *Russ. Chem. Bull.*, 2006, **55**, 369.
138 I. Ayala, J. J. P. Perry, J. Szczepanski, J. A. Tainer, M. T. Vala, H. S. Nick and D. N. Silverman, *Biophys. J.*, 2005, **89**, 4171.
139 X. Ren, C. Tu, D. Bhatt, J. J. P. Perry, J. A. Tainer, D. E. Cabelli and D. N. Silverman, *J. Mol. Struct.*, 2006, **790**, 168.
140 R. Gao, Y. Zhang, A. K. Choudhury, L. M. Dedkova and S. M. Hecht, *J. Am. Chem. Soc.*, 2005, **127**, 3321.
141 R. Gao, Y. Zhang, L. Dedkova, A. K. Choudhury, N. J. Rahier and S. M. Hecht, *Biochemistry*, 2006, **45**, 8402.
142 J. T. Stivers, S. Shuman and A. S. Mildvan, *Biochemistry*, 1994, **33**, 15449.
143 B. A. Diner and F. Rappaport, *Annu. Rev. Plant Biol.*, 2002, **53**, 551.
144 F. Rappaport, A. Boussac, D. A. Force, J. Peloquin, M. Brynda, M. Sugiura, S. Un, R. D. Britt and B. A. Diner, *J. Am. Chem. Soc.*, 2009, **131**, 4425.
145 D. Wang and P. A. Cole, *J. Am. Chem. Soc.*, 2001, **123**, 8883.
146 A. J. Ablooglu, J. H. Till, K. Kim, K. Parang, P. A. Cole, S. R. Hubbard and R. A. Kohanski, *J. Biol. Chem.*, 2000, **275**, 30394.
147 B. Gopishetty, L. Ren, T. M. Waller, A.-S. Wavreille, M. Lopez, A. Thakkar, J. Zhu and D. Pei, *Org. Lett.*, 2008, **10**, 4605.
148 M. Chalfie, Y. Tu, G. Euskirchen, W. W. Ward and D. C. Prasher, *Science*, 1994, **263**, 802.
149 V. V. Verkhusha and K. A. Lukyanov, *Nat. Biotechnol.*, 2004, **22**, 289.
150 A. Jiménez-Banzo, S. Nonell, J. Hofkens and C. Flors, *Biophys. J.*, 2008, **94**, 168.
151 J. H. Bae, P. Paramita Pal, L. Moroder, R. Huber and N. Budisa, *ChemBioChem*, 2004, **5**, 720.

152 P. P. Pal, J. H. Bae, M. K. Azim, P. Hess, R. Friedrich, R. Huber, L. Moroder and N. Budisa, *Biochemistry*, 2005, **44**, 3663.
153 S. Veettil, N. Budisa and G. Jung, *Biophys. Chem.*, 2008, **136**, 38.
154 B. J. Wilkins, S. Marionni, D. D. Young, J. Liu, Y. Wang, M. L. Di Salvo, A. Deiters and T. A. Cropp, *Biochemistry*, 2010, **49**, 1557.
155 G. Mayer and A. Heckel, *Angew. Chem., Int. Ed.*, 2006, **45**, 4900.
156 A. Goulding, S. Shrestha, K. Dria, E. Hunt and S. K. Deo, *Protein Eng., Des. Sel.*, 2008, **21**, 101.
157 Z. J. Xu, M. L. Love, L. Y. Ma, M. Blum, P. M. Bronskill, J. Bernstein, A. A. Grey, T. Hofmann, N. Camerman and J. T. Wong, *J. Biol. Chem.*, 1989, **264**, 4304.
158 C.-Y. Wong and M. R. Eftink, *Biochemistry*, 1998, **37**, 8947.
159 M. Lee and R. S. Phillips, *Bioorg. Med. Chem. Lett.*, 1991, **1**, 477.
160 M. Lee and R. S. Phillips, *Bioorg. Med. Chem.*, 1995, **3**, 195.
161 C. Giese, S. Lepthien, L. Metzner, M. Brandsch, N. Budisa and H. Lilie, *ChemMedChem*, 2008, **3**, 1449.
162 J. L. Hott and R. F. Borkman, *Biochem. J.*, 1989, **264**, 297.
163 P. M. Bronskill and J. T. Wong, *Biochem. J.*, 1988, **249**, 305.
164 A. Ozarowski, J. Q. Wu and A. H. Maki, *FEBS Lett.*, 1998, **422**, 52.
165 N. Budisa, C. Minks, F. J. Medrano, J. r. Lutz, R. Huber and L. Moroder, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 455.
166 C. Minks, R. Huber, L. Moroder and N. Budisa, *Biochemistry*, 1999, **38**, 10649.
167 D. A. Dougherty and D. A. Stauffer, *Science*, 1990, **250**, 1558.
168 Q.-s. Zhang, L. Shen, E.-d. Wang and Y.-l. Wang, *J. Protein Chem.*, 1999, **18**, 187.
169 N. Budisa, P. P. Pal, S. Alefelder, P. Birle, T. Krywcun, M. Rubini, W. Wenger, J. H. Bae and T. Steiner, *Biol. Chem.*, 2004, **385**, 191.
170 J. Panek, T. Ward, A. Jezierska and M. Novič, *J. Mol. Model.*, 2009, **15**, 257.
171 P. C. Weber, D. H. Ohlendorf, J. J. Wendoloski and F. R. Salemme, *Science*, 1989, **243**, 85.
172 M. Mae, H. Amii and K. Uneyama, *Tetrahedron Lett.*, 2000, **41**, 7893.
173 A. Vidal-Cros, M. Gaudry and A. Marquet, *Biochem. J.*, 1990, **266**, 749.
174 M. C. T. Hartman, K. Josephson and J. W. Szostak, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 4356.
175 M. C. T. Hartman, K. Josephson, C.-W. Lin and J. W. Szostak, *PLoS One*, 2007, **2**, e972.
176 H. J. C. Yeh, K. L. Kirk, L. A. Cohen and J. S. Cohen, *J. Chem. Soc., Chem. Commun.*, 1975, 928.
177 D. Y. Jackson, J. Burnier, C. Quan, M. Stanley, J. Tom and J. A. Wells, *Science*, 1994, **266**, 243.
178 D. S. Wimalasena, J. C. Cramer, B. E. Janowiak, S. J. Juris, R. A. Melnyk, D. E. Anderson, K. L. Kirk, R. J. Collier and J. G. Bann, *Biochemistry*, 2007, **46**, 14928.
179 D. S. Wimalasena, B. E. Janowiak, S. Lovell, M. Miyagi, J. Sun, H. Zhou, J. Hajduch, C. Pooput, K. L. Kirk, K. P. Battaile and J. G. Bann, *Biochemistry*, 2010, **49**, 6973.
180 J. F. Eichler, J. C. Cramer, K. L. Kirk and J. G. Bann, *ChemBioChem*, 2005, **6**, 2170.
181 G. Hortin, A. M. Stern, B. Miller, R. H. Abeles and I. Boime, *J. Biol. Chem.*, 1983, **258**, 4047.
182 M. A. Phillips, A. M. Stern, R. H. Abeles and A. H. Tashjian, *J. Pharmacol. Exp. Ther.*, 1983, **226**, 276.
183 P. K. Rathod, A. H. Tashjian and R. H. Abeles, *J. Biol. Chem.*, 1986, **261**, 6461.
184 T. Kitamoto, S. Marubayashi and T. Yamazaki, *Chem. Lett.*, 2006, 1264.
185 T. Kitamoto, S. Marubayashi and T. Yamazaki, *Tetrahedron*, 2008, **64**, 1888.
186 A. Yaron and F. Naider, *Crit. Rev. Biochem. Mol. Biol.*, 1993, **28**, 31.
187 G. Vanhoof, F. Goossens, I. De Meester, D. Hendriks and S. Scharpe, *FASEB J.*, 1995, **9**, 736.
188 F. Rabanal, M. D. Ludevid, M. Pons and E. Giralt, *Biopolymers*, 1993, **33**, 1019.
189 C. R. S. Briggs, D. O'Hagan, J. A. K. Howard and D. S. Yufit, *J. Fluorine Chem.*, 2003, **119**, 9.
190 N. Panasik, E. S. Eberhardt, A. S. Edison, D. R. Powell and R. T. Raines, *Int. J. Pept. Protein Res.*, 1994, **44**, 262.
191 M. L. DeRider, S. J. Wilkens, M. J. Waddell, L. E. Bretscher, F. Weinhold, R. T. Raines and J. L. Markley, *J. Am. Chem. Soc.*, 2002, **124**, 2497.
192 L. E. Bretscher, C. L. Jenkins, K. M. Taylor, M. L. DeRider and R. T. Raines, *J. Am. Chem. Soc.*, 2001, **123**, 777.
193 C. Renner, S. Alefelder, J. H. Bae, N. Budisa, R. Huber and L. Moroder, *Angew. Chem., Int. Ed.*, 2001, **40**, 923.
194 J. T. Gerig and R. S. McLeod, *J. Am. Chem. Soc.*, 1973, **95**, 5725.
195 J. A. Hodges and R. T. Raines, *J. Am. Chem. Soc.*, 2005, **127**, 15923.
196 W. Kim, K. I. Hardcastle and V. P. Conticello, *Angew. Chem., Int. Ed.*, 2006, **45**, 8141.
197 G.-Q. Shi and W.-L. Cai, *J. Org. Chem.*, 1995, **60**, 6289.
198 M. S. Weiss, A. Jabs and R. Hilgenfeld, *Nat. Struct. Mol. Biol.*, 1998, **5**, 676.
199 E. J. Milner-White, L. H. Bell and P. H. Maccallum, *J. Mol. Biol.*, 1992, **228**, 725.
200 E. S. Eberhardt, N. Panasik and R. T. Raines, *J. Am. Chem. Soc.*, 1996, **118**, 12261.
201 M. Tandon, M. Wu, T. P. Begley, J. Myllyharju, A. Pirskanen and K. Kivirikko, *Bioorg. Med. Chem. Lett.*, 1998, **8**, 1139.
202 K. L. Gorres, R. Edupuganti, G. R. Krow and R. T. Raines, *Biochemistry*, 2008, **47**, 9447.
203 A. A. Gottlieb, Y. Fujita, S. Udenfriend and B. Witkop, *Biochemistry*, 1965, **4**, 2507.
204 J. J. Hutton, A. Marglin, B. Witkop, J. Kurtz, A. Berger and S. Udenfriend, *Arch. Biochem. Biophys.*, 1968, **125**, 779.
205 K. L. Gorres and R. T. Raines, *Anal. Biochem.*, 2009, **386**, 181.
206 R. Golbik, C. Yu, E. Weyher-Stingl, R. Huber, L. Moroder, N. Budisa and C. Schiene-Fischer, *Biochemistry*, 2005, **44**, 16026.
207 T. T. Tran, N. Patino, R. Condom, T. Frogier and R. Guedj, *J. Fluorine Chem.*, 1997, **82**, 125.
208 N. B. Malkar, J. L. Lauer-Fields, J. A. Borgia and G. B. Fields, *Biochemistry*, 2002, **41**, 6054.
209 A. V. Persikov, J. A. M. Ramshaw, A. Kirkpatrick and B. Brodsky, *J. Am. Chem. Soc.*, 2003, **125**, 11500.
210 M. D. Shoulders, K. J. Kamer and R. T. Raines, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 3859.
211 C. Boulègue, A. G. Milbradt, C. Renner and L. Moroder, *J. Mol. Biol.*, 2006, **358**, 846.
212 T. Kitamoto, T. Ozawa, M. Abe, S. Marubayashi and T. Yamazaki, *J. Fluorine Chem.*, 2008, **129**, 286.
213 J. A. K. Howard, D. S. Yufit, O. V. Chetina, S. J. Teat, S. C. Capelli and P. Pattison, *Org. Biomol. Chem.*, 2010, **8**, 5110.
214 R. Bhattacharyya and P. Chakrabarti, *J. Mol. Biol.*, 2003, **331**, 925.
215 K. M. Thomas, D. Naduthambi and N. J. Zondlo, *J. Am. Chem. Soc.*, 2006, **128**, 2216.
216 T.-Y. Zheng, Y.-J. Lin and J.-C. Horng, *Biochemistry*, 2010, **49**, 4255.
217 C. J. McKnight, P. T. Matsudaira and P. S. Kim, *Nat. Struct. Mol. Biol.*, 1997, **4**, 180.
218 D. Naduthambi and N. J. Zondlo, *J. Am. Chem. Soc.*, 2006, **128**, 12430.
219 J. W. Neidigh, R. M. Fesinmeyer and N. H. Andersen, *Nat. Struct. Mol. Biol.*, 2002, **9**, 425.
220 C. Armishaw, A. A. Jensen, T. Balle, R. J. Clark, K. Harpsøe, C. Skonberg, T. Liljefors and K. Strømgaard, *J. Biol. Chem.*, 2009, **284**, 9498.
221 I. V. Maslennikov, Z. O. Shenkarev, M. N. Zhmak, V. T. Ivanov, C. Methfessel, V. I. Tsetlin and A. S. Arseniev, *FEBS Lett.*, 1999, **444**, 275.
222 S.-H. Hu, J. Gehrmann, L. W Guddat, P. F. Alewood, D. J. Craik and J. L. Martin, *Structure (London)*, 1996, **4**, 417.
223 D. Chow, M. L. Nunalee, D. W. Lim, A. J. Simnick and A. Chilkoti, *Mater. Sci. Eng., R.*, 2008, **62**, 125.
224 R. Parenteau-Bareil, R. Gauvin and F. Berthod, *Materials*, 2010, **3**, 1863.
225 S. K. Holmgren, K. M. Taylor, L. E. Bretscher and R. T. Raines, *Nature*, 1998, **392**, 666.
226 S. K. Holmgren, L. E. Bretscher, K. M. Taylor and R. T. Raines, *Chem. Biol.*, 1999, **6**, 63.
227 J. Bella, M. Eaton, B. Brodsky and H. M. Berman, *Science*, 1994, **266**, 75.

228 D. Barth, H. J. Musiol, M. Schütt, S. Fiori, A. G. Milbradt, C. Renner and L. Moroder, *Chem.–Eur. J.*, 2003, **9**, 3692.
229 J. A. Hodges and R. T. Raines, *J. Am. Chem. Soc.*, 2003, **125**, 9262.
230 M. Doi, Y. Nishi, S. Uchiyama, Y. Nishiuchi, T. Nakazawa, T. Ohkubo and Y. Kobayashi, *J. Am. Chem. Soc.*, 2003, **125**, 9922.
231 D. Barth, A. G. Milbradt, C. Renner and L. Moroder, *ChemBioChem*, 2004, **5**, 79.
232 M. Doi, Y. Nishi, S. Uchiyama, Y. Nishiuchi, H. Nishio, T. Nakazawa, T. Ohkubo and Y. Kobayashi, *J. Pept. Sci.*, 2005, **11**, 609.
233 R. Berisio, V. Granata, L. Vitagliano and A. Zagari, *J. Am. Chem. Soc.*, 2004, **126**, 11402.
234 A. Bondi, *J. Phys. Chem.*, 1964, **68**, 441.
235 Y. Nishi, M. Doi, S. Uchiyama, Y. Nishiuchi, T. Nakazawa, T. Ohkubo and Y. Kobayashi, *Lett. Pept. Sci.*, 2003, **10**, 533.
236 Y. Nishi, S. Uchiyama, M. Doi, Y. Nishiuchi, T. Nakazawa, T. Ohkubo and Y. Kobayashi, *Biochemistry*, 2005, **44**, 6034.
237 M. D. Shoulders and R. T. Raines, *Annu. Rev. Biochem.*, 2009, **78**, 929.
238 W. Kim, A. George, M. Evans and V. P. Conticello, *ChemBioChem*, 2004, **5**, 928.
239 W. Kim, R. A. McMillan, J. P. Snyder and V. P. Conticello, *J. Am. Chem. Soc.*, 2005, **127**, 18121.
240 W. J. Cook, H. Einspahr, T. L. Trapane, D. W. Urry and C. E. Bugg, *J. Am. Chem. Soc.*, 1980, **102**, 5502.
241 J. C. Horng and R. T. Raines, *Protein Sci.*, 2006, **15**, 74.
242 Y. C. Chiang, Y. J. Lin and J. C. Horng, *Protein Sci.*, 2009, **18**, 1967.
243 P. Ruzza, G. Siligardi, A. Donella-Deana, A. Calderan, R. Hussain, C. Rubini, L. Cesaro, A. Osler, A. Guiotto, L. A. Pinna and G. Borin, *J. Pept. Sci.*, 2006, **12**, 462.
244 T. Steiner, P. Hess, J. H. Bae, B. Wiltschi, L. Moroder and N. Budisa, *PLoS One*, 2008, **3**, e1680.
245 B. Holzberger and A. Marx, *J. Am. Chem. Soc.*, 2010, **132**, 15708.