

# Glycine homopeptides: the effect of the chain length on the crystal structure and solid state reactivity†

Cite this: DOI: 10.1039/c4ce00630e

Aaron J. Smith, Farukh I. Ali and Dmitriy V. Soldatov\*

A series of linear oligoglycines has been studied using X-ray diffraction, thermal analysis complemented with simultaneous FTIR spectroscopy, gas chromatography – mass spectrometry analysis, and  $^1\text{H}/^{13}\text{C}$  NMR. The new and previously reported data are rationalized to reveal the effect of the chain size on the crystal structure, molecular conformation and thermal stability of the oligopeptides in the solid state, as well as thermally induced transformations (pyrolysis) in the solid state and the gaseous phase. Tetraglycine (4G) and pentaglycine (5G) form triclinic crystals,  $P\bar{1}$ , where they adopt a zwitterionic form and fully extended chain conformation. The antiparallel  $\beta$  sheet arrangement is similar to that in  $\beta$ -triglycine ( $\beta$ 3G) and  $\alpha$ -diglycine ( $\alpha$ 2G) but differs from the Form I of polyglycine (polyG). The stability sequence  $\alpha$ 2G <  $\beta$ 3G < 4G < 5G has been attributed to the stabilizing environment in the solid phase. For all studied glycines, the thermal degradation is a kinetic process complicated by multiple parallel pathways and sensitivity to the experimental conditions. These pathways include defragmentation reactions (mainly deamination and decarboxylation) and condensation reactions (formation of new peptide bonds), the latter yielding cyclo-diglycine (c2G, or 2,5-diketopiperazine or DKP) as a major product. Further pyrolysis results in polymerization (solid samples) or further fragmentation (gaseous phase). The overall cyclization reaction  $\alpha$ 2G(solid) = c2G(solid) +  $\text{H}_2\text{O}$ (gas) is endothermic with  $\Delta H = 63.7(4)$  kJ mol $^{-1}$  at 492 K. The standard enthalpy of the reaction  $\alpha$ 2G(solid) = c2G(solid) +  $\text{H}_2\text{O}$ (liquid) was estimated to be  $\Delta H^{298} \approx 18$  kJ mol $^{-1}$  meaning the reaction is thermodynamically unfavorable under ambient conditions. Finally, the cyclization reaction mechanisms are discussed for the whole series studied.

 Received 27th March 2014,  
Accepted 3rd May 2014

DOI: 10.1039/c4ce00630e

[www.rsc.org/crystengcomm](http://www.rsc.org/crystengcomm)

## Introduction

A new found interest in short peptides has stemmed from the idea of utilizing these unique molecules in the design of template matrices for solid state reactions. Solid state organic synthesis (also known as “solvent-free synthesis”) has emerged as a new domain of synthetic organic chemistry<sup>1</sup> and to date, almost all main types of organic reactions have been conducted in the solid state.<sup>2,3</sup> The main advantages of such synthetic methods – simplicity and economy (solvent-free, often yield a single product without a need for separation), high (usually quantitative) yields, as well as specificity of the product (including stereospecificity) – resonate well with the principles of Green Chemistry.<sup>4</sup> However the placement of reactant

molecules in a desired mutual arrangement within the solid state is a challenging task preventing wide spread adaptation. One approach to resolve this issue is based on the co-crystallization of the reactant with suitable template molecules as demonstrated previously in some oligo- and polymerization reactions.<sup>5,6</sup> Peptides may be used to build these key template matrices as they readily form inclusion compounds and co-crystals with various organics.<sup>7</sup> Peptides are advantageous over other molecules due to the high degree of predictability of their crystal structure motifs and potential for modification simply by changing the side groups on the peptide backbone.<sup>8</sup> Short peptides are also non-toxic and ecofriendly and therefore the solids they form would be a suitable media to host “green” synthetic reactions.

The utilization of the peptide matrix in solid state synthesis requires a comprehensive knowledge of the molecular versus crystal structure relationship among peptides of different length and type, as well as the pathways of their solid state transformations induced by external stimuli, and the stability limits of the overall solids. These key solid state characteristics have not been sufficiently and systematically studied in their mutual relation when it comes to peptides.

Department of Chemistry, University of Guelph, 50 Stone Rd. E., Guelph, ON, N1G 2W1, Canada. E-mail: dsoldato@uoguelph.ca; Fax: +1 519 766 1499; Tel: +1 519 824 4120, ext 53548

† Electronic supplementary information (ESI) available: Crystal structure data in CIF format for 4G and 5G. Representative FTIR spectra of volatile products released during TGA experiments of  $\gamma$ 1G,  $\alpha$ 2G,  $\beta$ 3G, 4G, 5G and 2G(p-h). CCDC 993137 and 993138. For ESI and crystallographic data in CIF or other electronic format see DOI: 10.1039/c4ce00630e

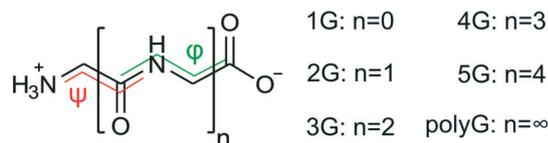


Fig. 1 Linear glycines in zwitterionic form: general formula and torsion angles describing molecular conformation.

For example, most studies on the solid state reactivity are limited to specific, pharmaceutically relevant peptides and proteins and pay minimal attention to their crystal structure.<sup>9–11</sup>

This work focuses on a series of linear glycine homopeptides (Fig. 1), their crystal structure, molecular conformation and solid state stability, as well as their thermally induced transformations in the solid *versus* gaseous phase. Since glycine is the simplest amino acid, oligoglycines are frequently used as models in experimental and theoretical studies on amide linkages and conformational preferences in peptides and proteins<sup>12–23</sup> as well as the degradation pathways of these important biomolecules.<sup>24–27</sup> Most of these studies refer to the molecules in vacuum or solution. Experimental solid state data are available only either for the smallest members of the series (1G, 2G and 3G) or the polypeptide (polyG). The medium-size oligomers are scarcely studied due to their low solubility and other limitations.<sup>28</sup> This study rationalizes, based on both newly obtained and previously reported data, several trends that occur in the series. In particular, we show how the chain length of the peptide affects the resulting crystal structure, as well as stability and thermally induced chemical transformations. Further, we compare the properties and behavior of the peptide molecules in the solid phase with the same molecules under other conditions.

## Experimental section

### General

All glycine oligomers were obtained from Chem-Impex (assays >99%) and used without further purification. The polymorphic forms of all previously studied glycines were determined using PXRD:  $\gamma$ 1G,<sup>29</sup>  $\alpha$ 2G<sup>30</sup> and  $\beta$ 3G.<sup>31</sup> These polymorphs are the most stable forms of the compounds at room temperature and atmospheric pressure.

Single crystals of 4G and 5G were obtained by crystallization from water.<sup>32</sup> The room temperature solubility of glycine oligomers in water decreases dramatically with the number of residues: 201.10, 72.35, 5.25 and 1.50 mg mL<sup>-1</sup> for 2G, 3G, 4G and 5G, respectively.<sup>33</sup> Due to this fact large volumes (>15 mL) of saturated solutions of 4G and 5G were prepared at ~90 °C and then were allowed to simultaneously evaporate and cool down to room temperature during two or three days, until <5% of the original volume remained. The crystals grew on the walls of the vial as extremely thin sheets aggregated into clusters. Samples suitable for XRD analysis were obtained after several (4G) or numerous (5G) crystallization attempts.

“Pre-heated” 2G, abbreviated here as “2G(p-h)”, was prepared by heating  $\alpha$ 2G (270 mg, 2.0 mmol) isothermally at 250 °C in a ventilation oven for 30 min. The mass loss of 18.3% was significantly higher than 13.6% calculated for the escape of water associated with the condensation reaction. The recovered sample was black in color due to a minor black fraction insoluble in water and DMSO. The soluble fraction revealed the presence of c2G (cyclo-diglycine) as the main product (>60% of the whole sample) as attested by <sup>1</sup>H NMR (the singlet of two NH protons at 8.04 ppm and doublet of four CH<sub>2</sub> protons at 3.72 ppm,  $J = 3.2$  Hz) and <sup>13</sup>C NMR (a signal from two C=O at 166.95 ppm and two CH<sub>2</sub> at 44.77 ppm). The presence of c2G as a dominant product was also confirmed by PXRD analysis of the recovered sample as its powder pattern corresponded to the crystal structure of c2G reported previously.<sup>34</sup> Finally, a crystal was selected from the sample (thin needle, 0.2 mm) and a single crystal XRD test (unit cell only) also confirmed c2G.<sup>34</sup> Another product of the soluble fraction (~25%) was not identified.<sup>35</sup> Heating 2G for 30 min at 280 °C resulted in similar results with a greater amount of insoluble fraction, while heating at 200 °C showed no reaction (PXRD pattern of  $\alpha$ 2G<sup>30</sup>); therefore only the product prepared at 250 °C was used in further experiments.

<sup>1</sup>H NMR and <sup>13</sup>C NMR spectra were collected on a Bruker Avance 400 MHz spectrometer. The samples (~2 mg) were dissolved in DMSO-d<sub>6</sub> (1 mL; D 99.9%, Cambridge Isotope Laboratories, Inc). The data were analyzed using the TopSpin 3.1 software.

### X-Ray diffraction (XRD)

All experiments were conducted at room temperature on a SuperNova Agilent single crystal diffractometer equipped with a microfocus CuK $\alpha$  ( $\lambda = 1.54184$  Å) radiation source and Atlas CCD detector. The data were processed using CrysAlisPro software.<sup>36</sup>

Powder X-ray diffraction (PXRD) experiments were performed with powder microsamples enclosed in 0.3 mm capillary tubes (Hampton Research) which were mounted on a goniometer head. Four frames were collected from different angular positions of the goniometer using  $\varphi$ -scan to cover  $2\theta$  range of 5–60°. Points of the powder pattern were generated from the original images with the step of 0.02° in  $2\theta$ , with each of the point representing readings from 200 to 2000 pixels.

Single crystals of 4G and 5G were mounted on the tip of a glass fiber. Diffraction intensity data were collected using  $\omega$ -scan to the maximum  $2\theta$  angle of 150.6° and 133.2° (resolution of 0.80 and 0.84 Å) for 4G and 5G, respectively, with the redundancy factor ~4 and completeness >99%. The unit cell parameters were refined using the entire data sets. The structures were solved (direct methods) and refined (full-matrix least-squares on  $F^2$ ) using SIR-92 (ref. 37) and SHELXL-97.<sup>38</sup> Non-hydrogen atoms were refined anisotropically, while hydrogen atoms were introduced at calculated positions and refined isotropically as riding, with  $U_{\text{iso}}$  values

set to  $1.2U_{eq}$  of the carrier atom. Three hydrogen atoms were located near each terminal nitrogen atom in a tetrahedral geometry implying the zwitterionic form in all cases. This form was also evident from the system of hydrogen bonds present. The distances and angles in the  $NH_3$  groups then were fixed but the groups were allowed to rotate around the N–C bond. Detailed analysis of the structures and preparation of graphics were carried out using WinGX<sup>39</sup> and Olex.<sup>40</sup> Crystal data and experimental parameters are summarized in Table 1. For further details see ESI.†

### Thermal analysis (DSC, TGA) and infra-red spectroscopy (FTIR)

The differential scanning calorimetry (DSC) and thermogravimetric analysis (TGA) measurements were conducted on a DSC Q-2000 and TGA Q-5000 IR analyzers, respectively (TA Instruments). The experiments were controlled using Q Advantage software and analyzed using Universal Analysis software associated with the instruments.

The DSC measurements were conducted with linear heating ( $5^\circ \text{ min}^{-1}$ ) in the 30–500 °C range under the flow of nitrogen ( $50 \text{ mL min}^{-1}$ ). The samples of  $\alpha 2G$  and  $2G(p-h)$  of various size (1.5 to 10.5 mg) were non-hermetically closed into Tzero aluminum pans. For the enthalpy measurements, the instrument was calibrated with indium before each experiment and two independent determinations were conducted, with 2.65 and 10.45 mg samples of  $\alpha 2G$ .

The TGA measurements were conducted with  $\sim 10$  mg samples placed in 100  $\mu\text{L}$  platinum pans and heated linearly from 30 to 800 °C in the flow of nitrogen. Two sets of the heating rate –  $N_2$  flow rate were used:  $5^\circ \text{ min}^{-1}$  –  $25 \text{ mL min}^{-1}$  and  $20^\circ \text{ min}^{-1}$  –  $100 \text{ mL min}^{-1}$ . The second set was used when the TGA measurements were accompanied by the FTIR analysis of the released volatiles. For this purpose, the TGA furnace outlet was connected to Thermo Scientific Nicolet iS10 FTIR

spectrometer through a stainless steel transfer line (2 mm internal diameter, coated with fused silica) heated to 220 °C and a TGA-FTIR accessory with 23 mL gas cell and 100 mm beam path length heated to 250 °C. The FTIR spectra were continuously collected in parallel with a TGA experiment at  $8 \text{ cm}^{-1}$  resolution, 8 scans per spectrum, and temporal resolution of 5.43 s. The FTIR experiments were controlled and analyzed using OMNIC Spectra software.

### Gas chromatography with mass spectrometry (GC-MS)

The samples were analyzed on an HP-5890 Series II gas chromatograph coupled to an HP-5971A mass selective detector. Hot aqueous solutions of the studied peptides were prepared immediately prior to the injection, with the concentration of  $2.5 \text{ mg mL}^{-1}$  for 5G and  $5 \text{ mg mL}^{-1}$  for all other compounds. About 8  $\mu\text{L}$  sample was injected in a split mode when injector was set to 300 °C (Restek 2 mm Gooseneck Splitless liner). A DB-5 ms column from Agilent was used (ID 0.25 mm, length 30 m, film depth 1  $\mu\text{m}$ ). The oven was programmed at 250 °C isothermally for 15.2 min. Purge valve was opened at 0.5 min at an  $8 \text{ mL min}^{-1}$  flow rate for 15 minutes, where helium (Grade 5, Linde) was used as a carrier gas. Solvent delay was set to 2.5 min. The detector, fitted with electron impact ionization source, was used at 300 °C. Total ion chromatograms were acquired scanning for 30–400 a.m.u. Compounds were identified and characterized by comparison of their retention times and mass spectra with mass spectral database (G1033A NIST/EPA/NIH-Library) associated with the instrument software.

## Results and discussion

### Crystal structures of 4G and 5G

Both 4G and 5G are triclinic, space group  $P\bar{1}$ . The centrosymmetry is possible due to the lack of a chiral center in glycine. In both structures, there are two crystallographically independent molecules (labeled A and B), both in a zwitterionic form and almost ideal fully extended chain conformation (Fig. 2 and 3; Table 2). The molecules assembly into a flat antiparallel  $\beta$  sheet banded arrangement, with  $N-H\cdots O=C$  hydrogen bonds between amide groups of the adjacent molecules (Fig. 2 and 3) as well as weak  $C_\alpha-H\cdots O=C$  hydrogen bonds. The  $\beta$  sheet bands are linked into infinite layers through charge-assisted  $^+N-H\cdots O^-$  hydrogen bonds on the edges of the bands, between the terminal amine and carboxyl groups. Also, the same type of hydrogen bonds cross-links the layers into a 3D framework. A more detailed comparison of the two crystal structures and the structures of other glycine homopeptides is provided in the following sections.

### Molecular conformations

The conformational behavior of glycine oligomers is of special interest. As the only unsubstituted amino acid, glycine is an important contributor to the conformational flexibility of peptide fragments in proteins. Collagen and silk fibroin

Table 1 Summary of crystallographic data

Compound	4G	5G
Empirical formula	$C_8H_{14}N_4O_5$	$C_{10}H_{17}N_5O_6$
<i>M</i>	246.2	303.3
Crystal system	Triclinic	Triclinic
Space group	$P\bar{1}$	$P\bar{1}$
<i>a</i> (Å)	4.8042(2)	4.7995(3)
<i>b</i> (Å)	14.8701(5)	14.9108(15)
<i>c</i> (Å)	15.6243(5)	17.6905(19)
$\alpha$ (°)	67.284(3)	90.203(8)
$\beta$ (°)	85.298(3)	94.866(7)
$\gamma$ (°)	88.190(3)	91.976(7)
<i>V</i> (Å <sup>3</sup> )	1026.14(7)	1260.7(2)
<i>Z</i>	4	4
<i>D</i> <sub>calc</sub> (g cm <sup>-3</sup> )	1.594	1.598
<i>T</i> (K)	295	295
Crystal color and habit	White sheet	White sheet
Crystal size (mm)	0.4 × 0.2 × 0.03	0.3 × 0.2 × 0.01
Reflections collected	19 837	17 005
Unique reflections ( <i>R</i> <sub>int</sub> )	4181 (0.030)	4425 (0.063)
Refined parameters	309	381
<i>R</i> <sub>1</sub> / <i>wR</i> <sub>2</sub> (data with <i>I</i> > 2 $\sigma$ <sub><i>i</i></sub> )	0.039/0.110	0.086/0.235
Residual min/max (e Å <sup>-3</sup> )	−0.23/+0.26	−0.31/+0.66

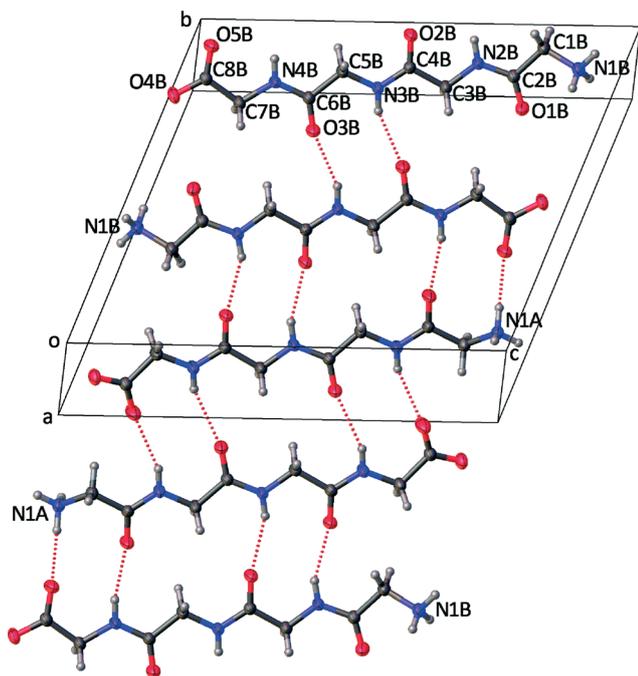


Fig. 2 Antiparallel  $\beta$  sheet arrangement of molecules in the crystal of 4G: ORTEP drawing and labeling scheme. The molecules in the sequence BBAA... form a flat band where the next neighbor is antiparallel and only the fourth neighbor is related by a translation.

are among many examples of highly flexible glycine-rich polymers occurring naturally. Conformationally flexible pentaglycine bridges cross-linking peptidoglycan layers in

*Staphylococcus aureus*' cell walls are responsible for the walls high stability, as well as surface protein attachment.<sup>43–45</sup> In the intact bacterium cells, the pentaglycine bridges predominantly adopt a compact, coil conformation<sup>46</sup> which easily rearranges into the extended conformation in response to a temperature change leading to swelling of the cell walls.<sup>47</sup> There is a slow exchange between the two conformations at 0 °C with the compact to extended form ratio of approximately 9 : 1.<sup>48</sup>

There are a number of computational studies in the literature comparing different conformations of oligoglycines and their analogs. For an isolated oligopeptide, some studies support the extended form as the most stable,<sup>18,41,49,50</sup> while others suggest a different conformation type.<sup>13,51</sup> Computations in solution environment do not support the extended conformation as energetically favorable.<sup>13,52</sup> Finally, experimental studies suggest the dominance of helical<sup>20,23,53</sup> or compact cyclic<sup>21</sup> forms in aqueous solutions.

In this context, it is very surprising that the 4G and 5G peptides in the crystal structures reported in this work adopt almost ideal fully extended conformation (Table 2). This is also true for the  $\beta$  polymorph of 3G.<sup>31</sup> There is a reasonably good agreement of the observed conformations with those calculated by Ivanova *et al.*<sup>41</sup> In contrast, the proposed conformation of polyG (form I) corresponds to a classical antiparallel  $\beta$  sheet (Table 2).<sup>42</sup> (Another polymorph of polyG, form II, contains a helical conformation.<sup>54</sup>) Moreover, the low molecular weight fraction of polyG may have a unique structure with the molecular conformation drastically

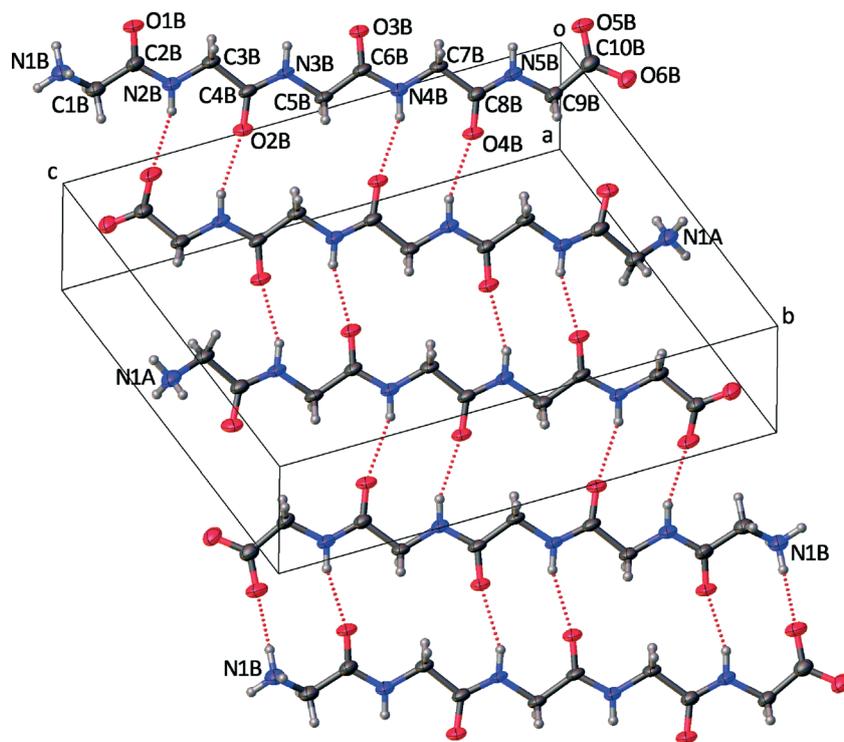


Fig. 3 Antiparallel  $\beta$  sheet arrangement of molecules in the crystal of 5G: ORTEP drawing and labeling scheme. The molecules in the sequence BBAA... form a flat band where the next neighbor is antiparallel and only the fourth neighbor is related by a translation.

**Table 2** Torsion angles ( $^{\circ}$ ) in some glycine homopeptides with a  $\beta$  sheet structure<sup>a,b</sup>

	$\varphi_{12}$	$\psi_{12}$	$\varphi_{23}$	$\psi_{23}$	$\varphi_{34}$	$\psi_{34}$	$\varphi_{45}$	$\psi_{45}$	Ref.
$\alpha$ 2G	-152.4	-155.1							Crystal <sup>30</sup>
2G	(145.3)	(156.8)							Single molecule, calculated <sup>c 41</sup>
$\beta$ 3G, mol. A	-149.9	+177.6	-171.5	+172.8					Crystal <sup>31</sup>
$\beta$ 3G, mol. B	-161.8	-165.9	+174.5	+172.8					Crystal <sup>31</sup>
3G	(148.5)	(177.7)	(172.8)	(173.1)					Single molecule, calculated <sup>c 41</sup>
4G, mol. A	-149.3	+178.1	-171.2	+173.7	-177.4	-172.3			Crystal, this work
4G, mol. B	-159.2	-165.2	+172.0	-175.6	+172.8	-174.6			Crystal, this work
4G	(169.2)	(170.4)	(175.3)	(178.9)	(170.6)	(176.1)			Single molecule, calculated <sup>c 41</sup>
5G, mol. A	-158.9	-163.9	+171.4	-175.0	+173.0	-176.0	+177.3	+174.2	Crystal, this work
5G, mol. B	+148.9	-178.7	+172.2	-172.6	+174.6	-175.1	+172.8	-173.7	Crystal, this work
I-polyG, chain 1	-149.9	+146.5							Crystal structure model <sup>42</sup>
I-polyG, chain 2	+149.9	-146.5							Crystal structure model <sup>42</sup>

<sup>a</sup>  $\varphi_{ij}$  and  $\psi_{ij}$  are the torsion angles (as defined in Fig. 1) between  $i$  and  $j$  residues. For example, for the molecule B in 4G (Fig. 2)  $\varphi_{12}$  is N1B-C1B-C2B-N2B and  $\psi_{12}$  is C2B-N2B-C3B-C4B. <sup>b</sup> Standard ( $\varphi$ ,  $\psi$ ) angle sets: (180,180) fully extended chain; (-142, 145) antiparallel-chain pleated  $\beta$  sheet; (-119, 113) parallel-chain pleated  $\beta$  sheet; (-47, -57) right-handed  $\alpha$  helix; (-49, -26) right-handed  $3_{10}$  helix. Simultaneous change of both signs produces a mirrored conformation. <sup>c</sup> Only absolute values reported.<sup>41</sup>

different from both fully extended and antiparallel  $\beta$  types.<sup>55</sup> These facts may suggest that although the same structural organization is observed in the  $\beta$ 3G – 4G – 5G series, these molecules are not long enough to observe the formation of a structure similar to that of polyG. A possible explanation for the difference between oligoglycines studied in this work and polyG is a high concentration of charge-bearing amine and carboxyl terminal groups in the former. These groups set additional requirements on the crystal structure due to the necessity of optimal charge distribution within the crystal as well as the formation of numerous energetically favorable charge-assisted hydrogen bonds.

### Solid state structure trends in glycine homopeptides

**Polymorphism.** 1G forms three polymorphs<sup>29</sup> of which  $\gamma$ 1G is the most thermodynamically stable at ambient conditions.<sup>56</sup> 2G forms three polymorphs,<sup>57</sup> of which  $\alpha$ 2G seems to be the most stable. No report of  $\gamma$ 2G has appeared since 1931 (ref. 57) and the full structure remains unknown. The crystal structure of  $\beta$ 2G<sup>58</sup> has H-bonded dimers and does

not resemble any of the classical protein structure types.  $\alpha$ 2G is the most stable polymorph at 105  $^{\circ}\text{C}$ <sup>58</sup> and, most likely, at ambient conditions. It has a monoclinic structure with one molecule in the asymmetric unit.<sup>30</sup> This structure is related to the family of antiparallel  $\beta$  sheet structures reported and discussed in this work (Table 3). 3G was isolated in two polymorphic forms<sup>59</sup> but no report of  $\alpha$ 3G has appeared since 1931 and the structure remains unknown.  $\beta$ 3G, which is apparently the stable form of 3G, crystallizes in a triclinic structure with two independent molecules (A and B, see Table 2). The structure is an antiparallel  $\beta$  sheet type. 4G and 5G crystallized and studied in this work always grew as a single polymorph (as attested by PXRD tests). Higher glycine oligomers, to the best of our knowledge, have not been characterized for their crystal structure. PolyG is known to exist as two polymorphs (forms I and II) of which form II is trigonal,  $P3_1$ , with the molecules having a helical conformation.<sup>54</sup> Form I is actually two polymorphs, both of the  $\beta$  sheet type.<sup>55</sup> The relative stability of the forms I and II is not known and would be difficult to determine experimentally due to the polymeric nature of the compound.

**Table 3** Comparison of the  $\beta$ -sheet-like structures of solid oligoglycines and polyG

	$\alpha$ 2G <sup>30</sup>	$\beta$ 3G <sup>31</sup>	4G <sup>a</sup>	5G <sup>a</sup>	PolyG <sup>b</sup>
Max van der Waals size of the molecule in the crystal ( $\text{\AA}$ )	9.9	13.5 (A); 13.6 (B)	17.2 (A and B)	20.8 (A); 20.7 (B)	3.52 per monomer
Repeat distance in the chain-axis direction ( $\text{\AA}$ )	9.42	24.07 (two molecules)	15.62 (=c)	17.69 (=c)	7.04 (=b; two monomers)
Order of molecules in the $\beta$ sheet band <sup>c</sup>	AA'...	AA'B'B...	AA'B'B...	AA'B'B...	AA'...
Direction of the $\beta$ sheet band in the crystal	[0 1 0] (along $b$ )	[0 1 $\bar{3}$ ]	[3 $\bar{1}$ 0]	[3 1 0]	[1 0 0] (along $a$ )
Average distance between adjacent molecules in the $\beta$ sheet band ( $\text{\AA}$ )	4.78	5.11	5.09	5.09	4.77 (=a/2)
Crystal planes defining the orientation of the $\beta$ sheet layers	(2 0 $\bar{2}$ )	(2 $\bar{3}$ $\bar{1}$ )	(1 3 0)	(1 $\bar{3}$ $\bar{2}$ )	(0 0 1)
Distance between adjacent $\beta$ sheet layers ( $\text{\AA}$ )	$d_{2\ 0\ \bar{2}} = 3.21^d$	$d_{2\ \bar{3}\ \bar{1}} = 3.29^e$	$d_{1\ \bar{3}\ 0} = 3.31^e$	$d_{1\ \bar{3}\ \bar{2}} = 3.32^e$	$d_{0\ 0\ 1} = 3.38^e$
Calculated density <sup>f</sup> ( $\text{g cm}^{-3}$ )	1.516	1.572	1.594	1.598	1.67

<sup>a</sup> Data of this work. <sup>b</sup> Form I; monoclinic cell ( $P2_1/c$ ) with  $a = 9.54$ ,  $b = 7.044$ ,  $c = 3.67$   $\text{\AA}$  and  $\beta = 113^{\circ}$ ; antiparallel rippled  $\beta$  sheet model used.<sup>42</sup> <sup>c</sup> A and B are two crystallographically independent molecules, while A and A' are enantiomorphs (related by an inversion center) and therefore antiparallel. <sup>d</sup> One of two most intense peaks on the simulated powder diffractogram. <sup>e</sup> The most intense peak on the simulated powder diffractogram. <sup>f</sup> Room temperature for all the structures.

**Trends in the  $\beta$  sheet structures.** The comparison of structures of the antiparallel  $\beta$  sheet type is presented in Table 3.  $\alpha$ 2G shows some resemblance to the higher oligomers and may be regarded as a simplified prototype in the series.  $\beta$ 3G, 4G and 5G have a remarkably similar crystal structure which may extend to the higher oligomers not studied yet. In all the three structures, there are two molecules (A and B) arranged in an antiparallel  $\beta$  sheet band as illustrated in Fig. 2 and 3. The band is very flat due to the fully extended molecular conformation of both A and B molecules and the H-bonds connecting neighboring molecules. The sequence of molecules in the band is AA'B'B..., where the two adjacent molecules of the same kind are related by an inversion center (e.g. A and A') while only the fourth neighbor is related to the initial by a translation. PolyG, form I, may exist in at least two variations<sup>42,55</sup> of which the most stable (for high molecular weight) has two enantiomorphic chains as H-bonded neighbors in the  $\beta$  sheet, resulting in the AA'... sequence.<sup>60</sup>

Table 3 reveals crystal packing similarities throughout the whole series, in spite of differences in the unit cell symmetry and dimensions. The maximum van der Waals size of the molecule changes by an increment of 3.5 to 3.7 Å per residue, as expected for an extended conformation. The low value of 3.5 Å for polyG is consistent with the chain conformation not being fully extended. The repeat distance in the chain-axis direction grows stepwise, with smaller steps from even to odd number of residues (2.62 Å from 2G to 3G; 2.07 Å from 4G to 5G) and greater otherwise (3.58 Å from 3G to 4G). The average distance between H-bonded molecules in the  $\beta$  sheet band falls into the narrow range of 4.8–5.1 Å defined by intermolecular H-bonding.

A very characteristic feature throughout the whole series is the distance between  $\beta$  sheet layers of 3.2–3.4 Å. It is defined by the interlayer van der Waals contacts (cf. C...C van der Waals contact of 3.4 Å). The distance accounts for the strongest (or second strongest) peak on the simulated powder diffractograms of the peptides which should be a signature in all  $\beta$  sheet glycines. For example, it appears as a strong peak at 3.40 Å<sup>42</sup> or 3.45 Å<sup>54</sup> on the powder diffractogram of polyG, form I, while there is no such a peak for polyG, form II (helical

structure).<sup>54</sup> The peak is absent for 3G·HCl<sup>61</sup> or c6G·0.5H<sub>2</sub>O (cyclo-hexaglycine hemihydrate)<sup>62</sup> as none of these structures have  $\beta$  sheet layers present. The powder patterns of antiparallel  $\beta$  sheet polyanalines all show the strongest peak corresponding to the interlayer distance of 5.1 Å, a substantial increase due to the methyl group of alanine.<sup>63</sup> One should note there are some variations within the narrow 3.2–3.4 Å range. As a consequence of not fully extended conformation in polyG, adjacent molecules in the  $\beta$  sheet band come slightly closer (4.77 Å) causing some increase in the interlayer distance (3.38 Å). As expected, the density of ideal structure grows with the number of residues, as the peptide covalent links effectively replace van der Waals contacts in the crystal.

### Stability trends in solid glycine homo peptides

Thermal stability of glycine oligomers under nitrogen atmosphere may be quantified using the onset temperatures of the initial mass loss on TGA curves (Fig. 4 and Table 4). As shown in the next section, the process associated with the initial mass loss is irreversible and therefore the onset temperatures are related to the kinetic stability of the compounds. Due to the kinetic nature of the process, the onset temperature depends on the heating rate; therefore only relative stability is discussed here.

One can clearly see that the stability limit increases with the chain length in the series of  $\beta$  sheet structures as  $\alpha$ 2G <  $\beta$ 3G < 4G < 5G. The main factor seems to be the stabilization environment in the solid phase. The peptide molecule must acquire a certain degree of freedom in order for any decomposition reaction to take place. Each additional residue adds two extra H-bonds as well as additional van der Waals contacts restricting the molecular motion. Another possible factor is the stability of the molecule itself. Although according to calculations the energy of the peptide bond does not depend significantly on the number of residues for a neutral oligoglycine in extended conformation,<sup>51</sup> in the studied solids the peptides exist in zwitterionic form, with two charges located on the opposite ends of the molecule. It is known that the zwitterionic form of amino acids and peptides is less stable for an

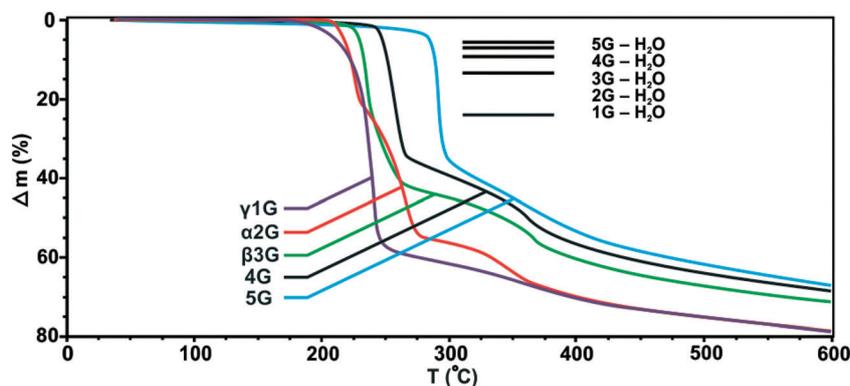


Fig. 4 TGA thermograms of studied glycines (stable polymorphs). The onset temperatures are listed in Table 4. Five bars on the right show calculated mass losses for the release of one mole of water for each of the five compounds. Purge: N<sub>2</sub>; heating rate: 5° min<sup>-1</sup>.

**Table 4** Onset temperatures (°C) of the initial mass loss of studied glycines subjected to TGA (average of three determinations)

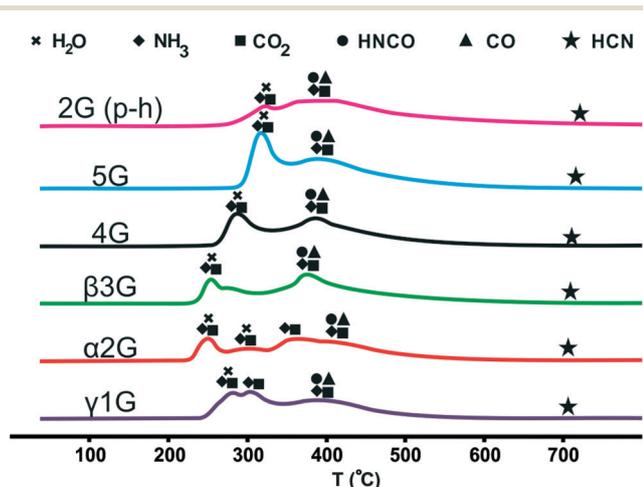
	$\gamma$ 1G	$\alpha$ 2G	$\beta$ 3G	4G	5G
Heating rate 5° min <sup>-1</sup>	221(2) <sup>a</sup>	216(1)	231(1)	250(1)	289(1)
Heating rate 20° min <sup>-1</sup>	244(2) <sup>a</sup>	226(1)	237(2)	262(2)	292(2)

<sup>a</sup> The value is derived from the tangent at the maximum rate mass loss but the gradual onset may indicate a delay due to an induction period (see text).

isolated molecule, but becomes stabilized in solution or the crystalline state. For example, from both experimental<sup>64,65</sup> and theoretical<sup>66–68</sup> studies glycine is not a zwitterion in the gaseous phase but is zwitterionic in any of its three crystalline forms.<sup>29</sup> Better separation of charges in the fully extended conformation of the studied oligoglycines may reduce the charge density and additionally stabilize longer-chain zwitterions. Finally, as shown in the next section, the principal decomposition pathways differ for the glycines of different chain length. In particular, the cyclization reaction for 1G cannot occur in a single step that may explain why the thermogram of  $\gamma$ 1G shows a more gradual onset, with the decomposition process delayed due to an induction period.

#### Pyrolysis of glycine homopeptides in the solid state

In order to examine the pathways of thermal decomposition under nitrogen (pyrolysis), the TGA experiments with simultaneous FTIR analysis of released volatiles were performed. The Gram–Schmidt curves showing total IR absorbance vs. temperature are shown in Fig. 5. Representative spectra are given in the ESI.† In all cases, the decomposition proceeds in several overlapped steps, with the same set of species produced: H<sub>2</sub>O, NH<sub>3</sub>, CO<sub>2</sub>, HNCO, CO and HCN. The formation of these volatiles has been observed and explained for 1G and 2G.<sup>27,69–74</sup> Samples recovered at 600 or 800 °C were black glassy residues which is consistent with the formation of partially decomposed polymeric matter.



**Fig. 5** Gram–Schmidt total IR absorbance vs. temperature for gaseous species evolved in TGA experiments (purge: N<sub>2</sub>, 100 mL min<sup>-1</sup>; heating rate: 20° min<sup>-1</sup>).

The pyrolysis of the studied compounds (Fig. 5) shows similarity in the sequence the volatile products are released although the temperatures of the major steps vary. At the decomposition onset, two types of processes occur. The first type is defragmentation reactions, specifically deamination and decarboxylation (*i.e.* loss of NH<sub>3</sub> and CO<sub>2</sub>) and possibly dehydrazination (the IR spectrum of hydrazine H<sub>2</sub>N–NH<sub>2</sub> is almost identical to that of NH<sub>3</sub>). These reactions should be followed by polymerization of remained less volatile fragments. The second type is the condensation reaction with the release of H<sub>2</sub>O and formation of new peptide bonds. The intermolecular condensation can lead to higher linear peptides while intramolecular condensation to cyclic oligopeptides of which the major product is the cyclo-diglycine, c2G, also known as 2,5-diketopiperazine (DKP).

At the temperatures above 350 °C the major volatiles released are NH<sub>3</sub>, CO<sub>2</sub>, HNCO and CO while H<sub>2</sub>O is not significant. The formation of NH<sub>3</sub> and CO<sub>2</sub> is still due to the defragmentation of linear peptides, through the stripping the amino and carboxyl terminals. Some of these terminals were present in the original molecules and survived the initial condensation step, while others could form in the reaction of the original molecules with c2G (*e.g.* 1G + c2G = 3G<sup>26,75</sup>). The condensation reaction is not noticeable here, probably because the concentration of pairs of amino and carboxyl terminals in close proximity to each other drops dramatically after the initial step. The release of HNCO and CO results from thermal decomposition of c2G (Fig. 6).<sup>27,76,77</sup>

Finally, at the highest temperatures some amounts of HCN are observed. This is also a product of c2G decomposition (Fig. 6). Although c2G produced at lower temperatures is too volatile to remain in the TGA crucible, some amounts of c2G can form through depolymerization reaction of higher oligopeptides<sup>24–27,77–80</sup> (see the last section). The formation of HCN over HNCO is caused by the higher temperature, as demonstrated in the literature.<sup>77</sup>

The differences in the decomposition temperature profiles along the series (Fig. 4 and 5) can be explained assuming the formation of c2G is the main process at the very beginning which in turn could accelerate other decomposition routes. The formation of c2G from 2G is the easiest as it requires the formation of only one peptide bond and the only other product is the H<sub>2</sub>O molecule. In contrast, two peptide bonds should form in order to yield c2G from 1G, and so the onset is the slowest in the series. For higher oligomers the onset temperature gradually increases with the chain length. This increase is due to the dependency of the depolymerization mechanism on the mobility of the molecules in the solid.

In order to confirm these conclusions, a sample of  $\alpha$ 2G was pre-heated in a ventilation oven at 250 °C and the product referred here as 2G(p-h) was characterized and further studied. The NMR and XRD analyses confirmed c2G as the main component of 2G(p-h) but the observed mass loss implied other reactions with a release of volatiles as well (see Experimental section for details). A sample of 2G(p-h) was subjected to a TGA-FTIR experiment under similar conditions

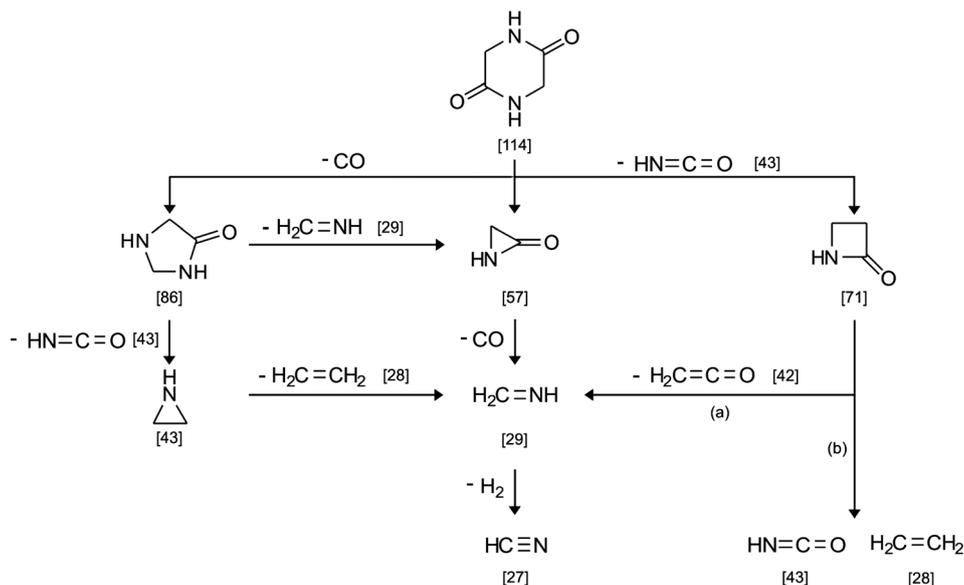


Fig. 6 Thermally induced decomposition pathways of c2G. The (a) and (b) defragmentation routes of  $\beta$ -lactam are the low- and high-temperature routes, respectively.<sup>77</sup> Molecular masses are shown in square brackets.

(Fig. 5). It is clear that although some condensation still occurs for 2G(p-h), it is unlikely to be the cyclization of 2G but rather depolymerization of longer chains, as the first peak is positioned similar to that of 5G.

### Pyrolysis in the gaseous phase and comparison

In order to compare the thermally induced transformations in the solid and gaseous phases, the same set of glycines was subjected to GC-MS analysis. At the injector site, a very small amount ( $\sim 8 \mu\text{L}$ ) of diluted peptide is instantly heated to  $300^\circ\text{C}$  allowing a reasonable assumption that all transformations induced at this temperature occur in the gaseous phase.

The gas chromatograms (Fig. 7) reveal only one peak. Although the intensity of the peak varies, the retention time and mass spectrum are practically identical in all cases. The mass spectra unequivocally indicate the formation of c2G from all the studied glycine homopeptides at  $300^\circ\text{C}$  in the gaseous phase. The absence of other peaks on the chromatograms is because other products that can form in the injector reaction space are either too light (e.g.  $\text{H}_2\text{O}$ ) or too heavy (e.g. higher oligomers), so that their retention times either too short or very long. In other words, c2G is the only product in the middle molecular mass range formed from the glycines gasified at  $300^\circ\text{C}$ . One could see that the intensity of the c2G peak is maximal for 2G(p-h), somewhat lower for 2G, 3G and 4G, relatively low for 5G and very low for 1G. This observation is consistent with the easiness of c2G formation from different molecules in the series. In particular, the 2G(p-h) sample already contains c2G, while the formation of c2G from 1G is difficult being more than a one-step reaction.

Fig. 6 shows the major decomposition pathways of the c2G molecule leading to species observed in the mass spectra. The three major products are  $\alpha$ -lactam,  $\beta$ -lactam and

4-imidazolidinone (3-, 4- and 5-membered rings respectively). The formation of these cyclic molecules and their decomposition produces smaller species:  $\text{HNCO}$ ,  $\text{H}_2\text{CNH}$ ,  $\text{H}_2\text{CCO}$ ,  $\text{H}_2\text{CCH}_2$ ,  $\text{CO}$ ,  $\text{H}_2$  and  $\text{HCN}$ . Some of these species appear on the FTIR spectra of volatiles produced in TGA of solid glycines (Fig. 5):  $\text{HNCO}$  and  $\text{CO}$  on moderate heating and  $\text{HCN}$  at high temperatures, while other species either do not form from the solid samples or quickly polymerize.

In general, one can see both similarities and differences in the pyrolysis of the glycine homopeptides in the solid state and gaseous phase. In both cases a number of parallel reactions occur, with TGA-FTIR and GS-MS being complementary in detecting these reactions and species produced. A similar set of final volatile products forms from the solid and gaseous glycines and c2G forms on early stages in both cases.

At the same time, there are evident differences that must be mentioned here. First, the pyrolysis of solid glycines produces polymeric materials due to high concentration of reactive species formed. In the gaseous phase, fragmentation is more likely. Consequently, a number of reactive species seen in the GC-MS experiments do not show in TGA-FTIR but polymeric residues are recovered instead. Second, the same reaction (such as the formation of c2G) in the solid state could be easier conducted and scaled. In the gaseous phase, only small amounts may be reacted and very big amounts of solvent and energy are required. These considerations bring forth the idea of using the solid state pyrolysis as a possible synthetic method for c2G and other DKPs as discussed in the next section.

### Formation of the cyclic dipeptide and cyclization mechanisms

The formation of cyclo-diglycine (c2G, or DKP) in the course of thermal treatment of the studied glycines is of special interest as it may suggest a general synthetic procedure for

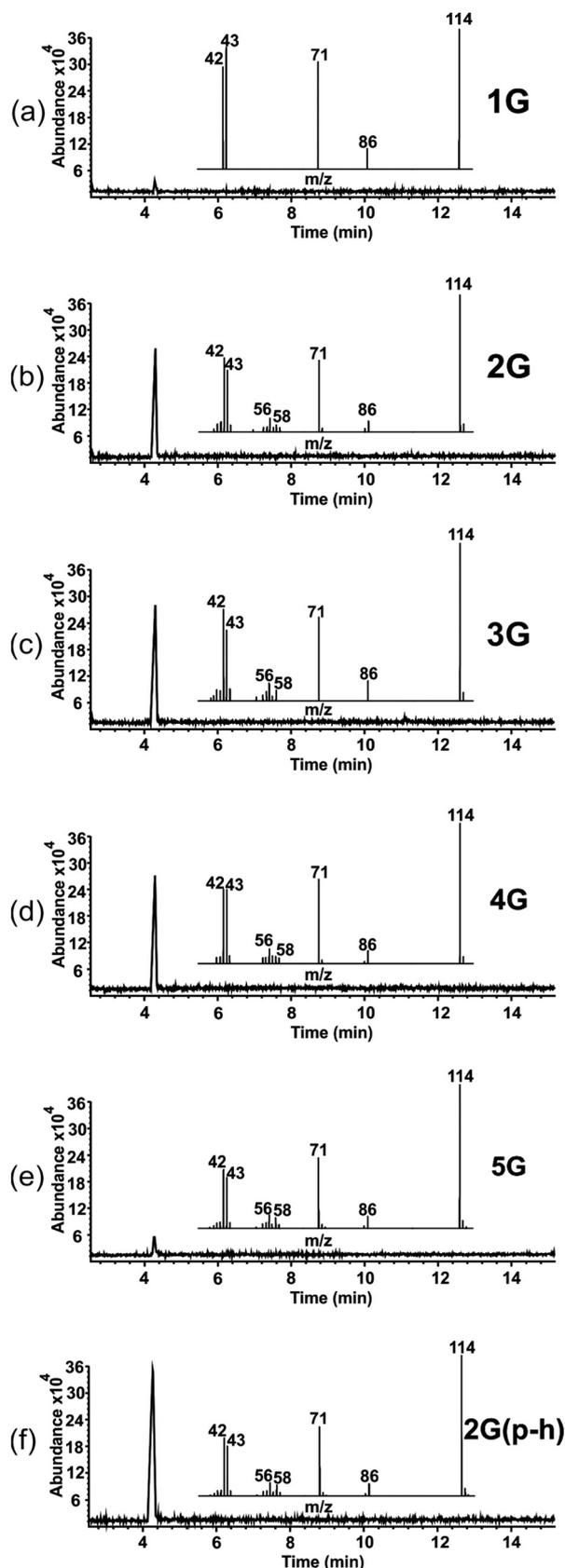


Fig. 7 Gas chromatograms of 1G–5G (a–e) and 2G (p–h) (f) and mass spectra corresponding to the major peak at ~4.3 min. The weight amounts of injected compounds are the same except for 5G which is half the amount (see Experimental section).

cyclic dipeptides (DKPs). These molecules attracted attention due to a wide spectrum of their biological activity and the use in medicinal chemistry.<sup>81–84</sup> On the other hand, the food industries interest in DKPs is due to their contribution to flavors, such as the bitter taste of roasted coffee,<sup>85</sup> meats,<sup>86</sup> beer<sup>87</sup> and chocolate.<sup>88</sup>

The formation of c2G, the simplest DKP, was observed in all thermal experiments of this work. However the reaction is almost always accompanied by various decomposition processes including the degradation of c2G itself. Therefore a better understanding of the reaction mechanism and thermodynamics are essential in the development of future solid state synthetic methods of DKPs.

**Thermodynamics.** The enthalpy of the cyclization reaction (1) was measured in DSC experiments:



$$\Delta H = 482(3) \text{ J g}^{-1} = 63.7(4) \text{ kJ mol}^{-1} \text{ (at 492 K)}$$

The DSC thermograms of  $\alpha 2G$  and 2G(p-h) are compared in Fig. 8. The only difference between the intact and pre-heated samples is the presence of an endotherm with the onset temperature of 219 °C due to the cyclization reaction (1). The complex endotherm at 280–330 °C, almost identical for both samples, corresponds to the sublimation of the cyclic form with evident partial degradation. The absence of any other effects on the  $\alpha 2G$  thermogram below 280 °C was surprising indicating the cyclization reaction was the only process in this range.<sup>89</sup> This observation made it possible to accurately measure the enthalpy of the reaction (1). The non-hermetic seal of the DSC capsules allowed for vaporization of water released during the reaction but prevented early sublimation by maintaining sufficient vapor pressure of c2G over the sample. Therefore the enthalpy of 63.7(4) kJ mol<sup>-1</sup> can be assigned to the pure reaction (1) at 219 °C. In order to estimate the standard enthalpy of the reaction, the vaporization enthalpy of water at 220 °C of 33.468 kJ mol<sup>-1</sup> (ref. 90) should be subtracted as well as the energy to heat water to 219 °C, 14.7 kJ mol<sup>-1</sup> (taking heat capacity of water 4.21 J g<sup>-1</sup> K<sup>-1</sup> = 75.9 J mol<sup>-1</sup> K<sup>-1</sup>).<sup>90</sup> The contribution due to heating the solids may be assumed to be -2.5 kJ mol<sup>-1</sup> (the heat capacities reported for 2G and c2G at 298 K are 149 (ref. 91) and 136 J mol<sup>-1</sup> K<sup>-1</sup>,<sup>92</sup> respectively). With this assumption, the standard enthalpy of the reaction (2) is



$$\Delta H^{298} \approx 18 \text{ kJ mol}^{-1}$$

As one can see, the cyclization reaction is thermodynamically unfavorable under ambient conditions (the entropy term is expected to be small) but becomes possible at elevated temperatures due to growing entropy term.

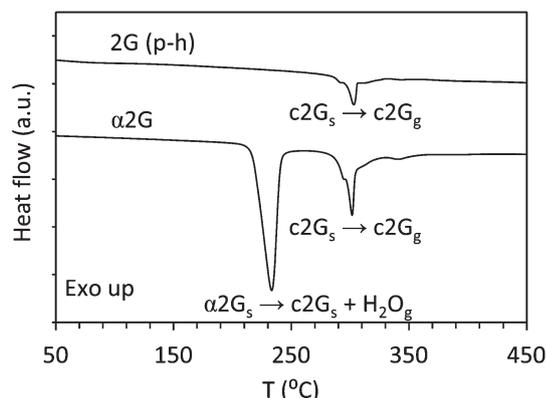


Fig. 8 DSC thermograms of 2G(p-h) and  $\alpha$ 2G (sample mass: 1.47 and 2.65 mg, respectively).

**Kinetics.** Aside from standard synthetic methods,<sup>93</sup> the formation of DKPs was observed in the mass spectrometry experiments on dipeptides,<sup>94–96</sup> pyrolyzed oligopeptides<sup>76</sup> and thermally degraded proteins,<sup>97</sup> upon sublimation of dipeptides,<sup>98,99</sup> from thermally<sup>100,101</sup> or microwave<sup>102–104</sup> treated solutions and in few cases in the solid state.<sup>105</sup> In all cases the process involves cyclization through intramolecular aminolysis that can proceed in different pathways. The type of the leaving group, amino acid residue, pH and temperature are the factors that greatly affect the mechanism and rate of the reaction in solution.<sup>100</sup> The most common mechanism observed for short peptides is a nucleophilic addition of the terminal amino group to the carbonyl moiety of the next residue in the peptide sequence. It outcomes in a tetrahedral intermediate followed by proton transfer from the amine nitrogen onto carbonyl oxygen. The resultant intermediate decomposes after the proton transfer to the leaving group.<sup>26,75,106,107</sup>

This mechanism is illustrated in Fig. 9 for the oligoglycines of this work. The tetrahedral intermediate is a cyclic zwitterion which splits into c2G and HR, where the leaving group HR is H<sub>2</sub>O for 2G or a shortened glycine chain for higher oligoglycines. The former case can be referred to as condensation, while the latter as depolymerization. Aside from the reactions in solution, the tetrahedral intermediate mechanism was proposed for mass spectrometry experiments,<sup>95</sup> pyrolysis of a melt,<sup>97</sup> or solid state pyrolysis of short peptides.<sup>78,79</sup> In contrast, the thermally-induced intramolecular cyclization of solid aspartame (Asp–Phe–OCH<sub>3</sub>) follows the S<sub>N</sub>1 mechanism. The first step involves the release of OCH<sub>3</sub><sup>−</sup> leaving group to form acyl cation, followed by the attack of the

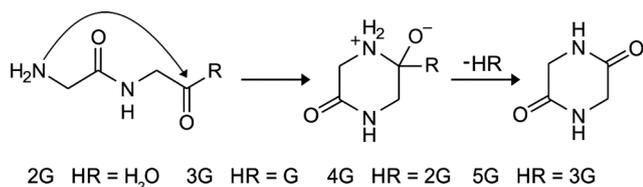


Fig. 9 Thermally induced intramolecular cyclization of oligoglycines through a tetrahedral intermediate.

terminal amino group on the carbonyl to close the ring.<sup>108</sup> Finally other, less common mechanisms were proposed.<sup>109</sup> One reason why the tetrahedral intermediate mechanism (Fig. 9) is believed to take place in this study is because it agrees with all experimental observations and explains the similarity of the decomposition process for the whole series. Another reason is that the hydroxyl ion is not a good leaving group making the S<sub>N</sub>1 mechanism very unlikely.

As described in the crystal structure part, the peptide solids of this work have their molecules in the zwitterionic form and fully extended conformation packed in a  $\beta$  sheet arrangement. For intramolecular aminolysis, the peptide molecule should be in its canonical rather than zwitterionic form. Further complications include restricted motion of the molecule and the necessity to break a number of intermolecular hydrogen bonds as the reaction proceeds in the solid state. These considerations imply a multi-step process much more complicated than would appear from Fig. 9. The thermal energy appears to be essential in generating a neutral peptide molecule, breaking hydrogen bonds and disrupting the crystal packing in order for the molecule to fold and close the cycle. Although more detailed studies are needed to elucidate the mechanism further, it is relevant to mention here a remarkable flexibility observed previously in crystalline peptides.<sup>110–114</sup> Significant local distortions of the crystal structure and possibly concerted mechanisms may be expected. The estimated activation energies for cyclization of aspartame and aspartyl-phenylalanine of 268(8) and 242(8) kJ mol<sup>−1</sup> respectively<sup>115</sup> may illustrate the kinetic requirements for the reactions to occur in the solid state, the values exceeding the cyclization enthalpy determined in this work by about one order of magnitude.

## Conclusion

This work reveals the effect of the chain size on the properties of solid peptides made of glycine residues, including their crystal structure, conformation, spatial arrangement and non-valent interactions of the peptide molecules in the crystal, as well as thermal stability and chemical changes in the solid state induced by heating.

The antiparallel  $\beta$  sheet structure with the molecules in zwitterionic form and fully extended conformation is the most stable crystalline form for lower glycine oligomers although up to three polymorphs may be isolated. For polyglycine the crystal structure and the molecular conformation differ, but one of the polymorphs (form I), likely the most stable, is still a  $\beta$  sheet structure. In spite of dissimilar unit cell parameters, the crystal packing motifs are very similar throughout the series.

The thermal stability of crystalline oligoglycines increases with the chain length due to the stabilization environment of the molecules in the crystal and other factors. The decomposition onset is subtle and involves a number of possible pathways which depend, to a lesser extent, on the peptide chain length but, to a much greater extent, on the experimental conditions. The two types of reactions at the onset are

defragmentation and condensation. These two reactions take place both in the solid state and the gas phase but polymerization of fragments dominates on a later stage for solid state samples only.

A major condensation product in thermal treatment of all oligoglycines is cyclo-diglycine (DKP). This formation occurs as intramolecular cyclization from diglycine and a depolymerization reaction from higher oligomers. From diglycine, the reaction proceeds quickly and DKP may be obtained as a pure product but the reaction is too sensitive to experimental conditions and difficult to scale. The solid state synthesis of DKP could be a synthetically important reaction if conducted without side products. More studies are needed in order to understand the effect of various factors on the cyclization reaction and to prevent undesired degradation processes which occur in parallel. In particular, the mechanism of the cyclization in the solid state appears to be complex and needs to be understood. The outcome of such studies could be the development of a general synthetic procedure for various DKPs.

Finally, similar studies on other peptide series would be useful to verify how the trends observed in our study apply to solid peptides of various compositions.

## Acknowledgements

We thank the Canada Foundation for Innovation (CFI), Ontario Ministry of Research and Innovation (MRI) and Department of Chemistry at the University of Guelph for funding the X-ray diffraction facility, and The National Science and Engineering Research Council Canada for the thermoanalytical instrumentation. DVS is grateful to Prof. W. Gabryelski (U. of Guelph) for a useful discussion. The assistance of M. Ignacio in some of the PXRD tests is appreciated.

## Notes and references

- G. M. J. Schmidt, *Pure Appl. Chem.*, 1971, **27**, 647–678.
- K. Tanaka and F. Toda, *Chem. Rev.*, 2000, **100**, 1025–1074.
- G. Kaupp, *Top. Curr. Chem.*, 2005, **254**, 95–183.
- P. T. Anastas and J. C. Warner, *Green chemistry: theory and practice*, Oxford University Press, New York, 1998.
- L. R. MacGillivray, G. S. Papaefstathiou, T. Friščić, D. B. Varshney and T. D. Hamilton, *Top. Curr. Chem.*, 2004, **248**, 201–221.
- J. W. Lauher, F. W. Fowler and N. S. Goroff, *Acc. Chem. Res.*, 2008, **41**, 1215–1229.
- D. V. Soldatov, in *Nanoporous Materials*, ed. A. Sayari and M. Jaroniec, World Scientific, New Jersey, 2008, pp. 213–224.
- C. H. Görbitz, *Acta Crystallogr., Sect. B: Struct. Sci.*, 2010, **66**, 84–93.
- M. C. Lai and E. M. Topp, *J. Pharm. Sci.*, 1999, **88**, 489–499.
- S. Jacob, A. A. Shirwaikar, K. K. Srinivasan, J. Alex, S. L. Prabu, R. Mahalaxmi and R. Kumar, *Indian J. Pharm. Sci.*, 2006, **68**, 154–163.
- J. Patel, R. Kothari, R. Tunga, N. M. Ritter and B. S. Tunga, *BioProcess Int.*, 2011, **9**(1), 20–31.
- R. A. Scott and H. A. Scheraga, *J. Chem. Phys.*, 1966, **45**, 2091–2101.
- C. Alemán and J. J. Perez, *J. Comput.-Aided Mol. Des.*, 1993, **7**, 241–250.
- T. V. Chalikian, A. P. Sarvazyan, T. Funck and K. J. Breslauer, *Biopolymers*, 1994, **34**, 541–553.
- C. Alemán, R. Roca, F. J. Luque and M. Orozco, *Proteins*, 1997, **28**, 83–93.
- R. Improta, V. Barone, K. N. Kudin and G. E. Scuseria, *J. Chem. Phys.*, 2001, **114**, 2541–2549.
- G. Nandini and D. N. Sathyanarayana, *J. Phys. Chem. A*, 2003, **107**, 11391–11400.
- A. Chatterjee, L. Zhao, L. Zhang, D. Pradhan, X. Zhou and K. T. Leung, *J. Chem. Phys.*, 2008, **129**, 1–6.
- R. Wu and T. B. McMahon, *J. Phys. Chem. B*, 2009, **113**, 8767–8775.
- S. Bykov and S. Asher, *J. Phys. Chem. B*, 2010, **114**, 6636–6641.
- S. A. Allison, H. Pei, U. Twahir, H. Wu and H. Cottet, *J. Sep. Sci.*, 2010, **33**, 2430–2438.
- A. M. Ferrari, B. Civalleri and R. Dovesi, *J. Comput. Chem.*, 2010, **31**, 1777–1784.
- S. V. Bykov and S. A. Asher, *J. Phys. Chem. Lett.*, 2010, **1**, 269–271.
- Y.-C. Oh, C.-K. Shu and C.-T. Ho, *J. Agric. Food Chem.*, 1991, **39**, 1553–1554.
- V. A. Basiuk and J. Douda, *J. Anal. Appl. Pyrolysis*, 2000, **55**, 235–246.
- J. Bujdák and M. Rode, *J. Pept. Sci.*, 2004, **10**, 731–737.
- J. Li, Z. Wang, X. Yang, L. Hu, Y. Liu and C. Wang, *J. Anal. Appl. Pyrolysis*, 2007, **80**, 247–253.
- The oligomers dissolve only in highly polar solvents, easily become ionic in acidic or basic media and gradually degrade with time.
- E. V. Boldyreva, T. N. Drebushchak and E. S. Shutova, *Z. Kristallogr.*, 2003, **218**, 366–376.
- S. A. Moggach, D. R. Allan, S. Parsons and L. Sawyer, *Acta Crystallogr., Sect. B: Struct. Sci.*, 2006, **62**, 310–320.
- T. Srikrishnan, N. Winiewicz and R. Parthasarathy, *Int. J. Pept. Protein Res.*, 1982, **19**, 103–113.
- The solubility in other solvents, such as methanol, ethanol or their mixtures with water, was too low.
- J. Lu, X.-J. Wang, X. Yang and C.-B. Ching, *J. Chem. Eng. Data*, 2006, **51**, 1593–1596.
- R. Degeilh and R. E. Marsh, *Acta Crystallogr.*, 1959, **12**, 1007–1014.
- <sup>1</sup>H NMR: triplet at 1.19 ppm, singlet at 2.01 ppm and quartet at 4.04 ppm (approx. intensity ratio 2:1.5:1). PXRD:  $2\theta = 20.46, 21.28$  and  $25.85^\circ$ .
- Xcalibur CCD System, CrysAlisPro Software, Version 1.171.35.8*, Agilent Technologies, Yarnton, Oxfordshire, UK, 2011.
- A. Altomare, G. Cascarano, C. Giacovazzo and A. Gualardi, *J. Appl. Crystallogr.*, 1993, **26**, 343–350.
- G. M. Sheldrick, *SHELXL-97, Program for refinement of crystal structures*, University of Göttingen, Germany, 1997.
- L. J. Farrugia, *J. Appl. Crystallogr.*, 1999, **32**, 837–838.

- 40 O. V. Dolomanov, L. J. Bourhis, R. J. Gildea, J. A. K. Howard and H. Puschmann, *J. Appl. Crystallogr.*, 2009, **42**, 339–341.
- 41 B. B. Ivanova, T. Kolev and S. Y. Zareva, *Biopolymers*, 2006, **82**, 587–596.
- 42 B. Lotz, *J. Mol. Biol.*, 1974, **87**, 169–180.
- 43 O. Schneewind, A. Fowler and K. F. Faull, *Science*, 1995, **268**, 103–106.
- 44 X. Zhou and L. Cegelski, *Biochemistry*, 2012, **51**, 8143–8153.
- 45 A. P. A. Hendrickx, J. M. Budzik, S.-Y. Oh and O. Schneewind, *Nat. Rev. Microbiol.*, 2011, **9**, 166–176.
- 46 A. Lapidot and C. S. Irving, *Proc. Natl. Acad. Sci. U. S. A.*, 1977, **74**, 1988–1992.
- 47 A. Lapidot and C. S. Irving, *Biochemistry*, 1979, **9**, 1788–1796.
- 48 G. Tong, Y. Pan, H. Dong, R. Pryor, G. E. Wilson and J. Schaefer, *Biochemistry*, 1997, **36**, 9859–9866.
- 49 F. A. Momany, R. F. McGuire, J. F. Yan and H. A. Scheraga, *J. Phys. Chem.*, 1971, **75**, 2286–2297.
- 50 L. L. Shipman and R. E. Christoffersen, *Proc. Natl. Acad. Sci. U. S. A.*, 1972, **69**, 3301–3304.
- 51 P. Chaudhuri and S. Canuto, *J. Mol. Struct.: THEOCHEM*, 2008, **849**, 25–32.
- 52 R. Importa, V. Barone, K. N. Kudin and G. E. Scuseria, *J. Chem. Phys.*, 2001, **114**, 2541–2549.
- 53 S. Ohnishi, H. Kamikubo, M. Onitsuka, M. Kataoka and D. Shortle, *J. Am. Chem. Soc.*, 2006, **128**, 16338–16344.
- 54 F. H. C. Crick and A. Rich, *Nature*, 1955, **176**, 780–781.
- 55 A. V. Kajava, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 1999, **55**, 436–442.
- 56 E. V. Boldyreva, V. A. Drebuschak, T. N. Drebuschak, I. E. Paukov, Y. A. Kovalevskaya and E. S. Shutova, *J. Therm. Anal. Calorim.*, 2003, **73**, 409–418.
- 57 J. D. Bernal, *Z. Kristallogr., Kristallgeom., Kristallphys., Kristallchem.*, 1931, **78**, 363–369.
- 58 E. W. Hughes and W. J. Moore, *J. Am. Chem. Soc.*, 1949, **71**, 2618–2623.
- 59 F. V. Lenel, *Z. Kristallogr., Kristallgeom., Kristallphys., Kristallchem.*, 1931, **81**, 224–229.
- 60 F. Colonna-Cesari, S. Premilat and B. Lotz, *J. Mol. Biol.*, 1974, **87**, 181–191.
- 61 V. Lalitha and E. Subramanian, *Cryst. Struct. Commun.*, 1982, **11**, 561–564.
- 62 I. L. Karle and J. Karle, *Acta Crystallogr.*, 1963, **16**, 969–975.
- 63 T. Asakura, M. Okonogi, K. Horiguchi, A. Aoki, H. Saitô, D. P. Knight and M. P. Williamson, *Angew. Chem., Int. Ed.*, 2012, **51**, 1212–1215.
- 64 R. D. Suenram and F. J. Lovas, *J. Mol. Spectrosc.*, 1980, **72**, 372–382.
- 65 M. J. Locke and R. T. McIver Jr., *J. Am. Chem. Soc.*, 1983, **105**, 4226–4232.
- 66 Y. Ding and K. Krogh-Jespersen, *Chem. Phys. Lett.*, 1992, **199**, 261–266.
- 67 D. Yu, D. A. Armstrong and A. Rauk, *Can. J. Chem.*, 1992, **70**, 1762–1772.
- 68 J. H. Jensen and M. S. Gordon, *J. Am. Chem. Soc.*, 1995, **117**, 8159–8170.
- 69 P. G. Simmonds, E. E. Medley, M. A. Ratcliff and G. P. Shulman, *Anal. Chem.*, 1972, **44**, 2060–2066.
- 70 K. J. Voorhees, F. Basile, M. B. Beverly, C. Abbas-Hawks, A. Hendricker, R. B. Cody and T. L. Hadfield, *J. Anal. Appl. Pyrolysis*, 1997, **40–41**, 111–134.
- 71 N. Gallois, J. Templier and S. Derenne, *J. Anal. Appl. Pyrolysis*, 2007, **80**, 216–230.
- 72 V. Y. Yablokov, I. L. Smel'tsova, I. A. Zelyaev and S. V. Mitrofanova, *Russ. J. Gen. Chem.*, 2009, **79**, 1704–1706.
- 73 S.-S. Choi and J.-E. Ko, *J. Anal. Appl. Pyrolysis*, 2010, **89**, 74–86.
- 74 S. C. Moldoveanu, *Pyrolysis of Organic Molecules with Applications to Health and Environmental Issues*, in *Techniques and Instrumentation in Analytical Chemistry*, Elsevier, Amsterdam, 2010, ch. 18, vol. 28.
- 75 N. Nagayama, O. Takaoka, K. Inomata and Y. Yamagata, *Origins Life Evol. Biosphere*, 1990, **20**, 249–257.
- 76 G. Chiavari and G. C. Galletti, *J. Anal. Appl. Pyrolysis*, 1992, **24**, 123–137.
- 77 K.-M. Hansson, J. Samuelsson, L.-E. Åmand and C. Tullin, *Fuel*, 2003, **82**, 2163–2172.
- 78 K. J. Voorhees, W. Zhang, A. D. Hendricker and B. Murugaverl, *J. Anal. Appl. Pyrolysis*, 1994, **30**, 1–16.
- 79 K.-M. Hansson, L.-E. Åmand, A. Habermann and F. Winter, *Fuel*, 2003, **82**, 653–660.
- 80 K.-M. Hansson, J. Samuelsson, C. Tullin and L.-E. Åmand, *Combust. Flame*, 2004, **137**, 265–277.
- 81 C. Prasad, *Peptides*, 1995, **16**, 151–164.
- 82 M. B. Martins and I. Carvalho, *Tetrahedron*, 2007, **63**, 9923–9932.
- 83 A. D. Borthwick, *Chem. Rev.*, 2012, **112**, 3641–3716.
- 84 Y. Wang, P. Wang, H. Ma and W. Zhu, *Expert Opin. Ther. Pat.*, 2013, **23**, 1415–1433.
- 85 M. Ginz and U. H. Engelhardt, *J. Agric. Food Chem.*, 2000, **48**, 3528–3532.
- 86 M. Z. Chen, M. L. Dewis, K. Kraut, D. Merritt, L. Reiber, L. Trinnaman and N. C. Da Costa, *J. Food Sci.*, 2009, **74**, C100–C105.
- 87 M. Gautschi, J. P. Schmid, T. L. Peppard, T. P. Ryan, R. M. Tuorto and X. Yang, *J. Agric. Food Chem.*, 1997, **45**, 3183–3189.
- 88 T. Stark and T. Hofmann, *J. Agric. Food Chem.*, 2005, **53**, 7222–7231.
- 89 As shown in this study, a number of reactions may be induced upon thermal treatment of oligoglycines. In most experiments cyclization is accompanied by side degradation reactions which are difficult to prevent. The observation of cyclization as a single process in DSC experiments may be attributed to the small size and purity of the samples, as well as the dynamic nature of this technique.
- 90 *CRC Handbook of Chemistry and Physics*, ed. D. R. Lide, CRC Press, Boca Raton, 88th edn, 2007–2008, p. 6–4.
- 91 V. G. Badelin, O. V. Kulikov, V. S. Vatagin, E. Udzig, A. Zielenkiewicz, W. Zielenkiewicz and G. A. Krestov, *Thermochim. Acta*, 1990, **169**, 81–93.
- 92 G. Barone and R. Puliti, *J. Therm. Anal. Calorim.*, 1999, **57**, 119–132.

- 93 E. Giralt, R. Eritja, J. Josa, C. Kuklinski and E. Pedroso, *Synthesis*, 1985, 181–184.
- 94 H. J. Svec and G. A. Junk, *J. Am. Chem. Soc.*, 1964, **86**, 2278–2282.
- 95 I. Lüderwald, M. Przybylski, H. Ringsdorf and D. Silberhorn, *Z. Naturforsch., B: Anorg. Chem., Org. Chem.*, 1978, **33**, 805–808.
- 96 A. S. Noguerola, B. Murugaverl and K. J. Voorhees, *J. Am. Soc. Mass Spectrom.*, 1992, **3**, 750–756.
- 97 M. Langhammer, I. Lüderwald and A. Simons, *Fresenius' Z. Anal. Chem.*, 1986, **324**, 5–8.
- 98 D. Gross and G. Grodsky, *J. Am. Chem. Soc.*, 1955, **77**, 1678–1680.
- 99 V. A. Basiuk and T. Y. Gromovoy, *Collect. Czech. Chem. Commun.*, 1994, **59**, 461–466.
- 100 P. M. Fisher, *J. Pept. Sci.*, 2003, **9**, 9–35.
- 101 P. Gomez, N. Vale and R. Moreira, *Molecules*, 2007, **12**, 2484–2506.
- 102 L. Pérez-Picaso, J. Escalante, H. F. Olivo and M. Y. Rios, *Molecules*, 2009, **14**, 2836–2849.
- 103 M. Jainta, M. Nieger and S. Bräse, *Eur. J. Org. Chem.*, 2008, 5418–5424.
- 104 Nonappa, K. Ahonen, M. Lahtinen and E. Kolehmainen, *Green Chem.*, 2011, **13**, 1203–1209.
- 105 S.-Y. Lin and S.-L. Wang, *Adv. Drug Delivery Rev.*, 2012, **64**, 461–478.
- 106 N. F. Sepetov, M. A. Krymsky, M. V. Ovchinnikov, Z. D. Bepalova, O. L. Isakova, M. Soucek and M. Lebl, *Pept. Res.*, 1991, **4**, 308–313.
- 107 P. Xia, C. Wang and C. Qi, *Chin. J. Chem.*, 2013, **31**, 813–818.
- 108 Y.-D. Cheng and S.-Y. Lin, *J. Agric. Food Chem.*, 2000, **48**, 631–635.
- 109 J. Templier, N. Gallois and S. Derenne, *J. Anal. Appl. Pyrolysis*, 2013, **104**, 684–694.
- 110 C. H. Görbitz, *Acta Crystallogr., Sect. B: Struct. Sci.*, 2002, **58**, 849–854.
- 111 M. Akazome, A. Hirabayashi, K. Takaoka, S. Nomura and K. Ogura, *Tetrahedron*, 2005, **61**, 1107–1113.
- 112 D. V. Soldatov, I. L. Moudrakovski, E. V. Grachev and J. A. Ripmeester, *J. Am. Chem. Soc.*, 2006, **128**, 6737–6744.
- 113 T. J. Burchell, D. V. Soldatov, G. D. Enright and J. A. Ripmeester, *CrystEngComm*, 2007, **9**, 922–929.
- 114 R. Anedda, D. V. Soldatov, I. L. Moudrakovski, M. Casu and J. A. Ripmeester, *Chem. Mater.*, 2008, **20**, 2908–2920.
- 115 S. S. Leung and D. J. W. Grant, *J. Pharm. Sci.*, 1997, **86**, 64–71.