

ARTICLE

<https://doi.org/10.1038/s42004-019-0224-2>

OPEN

How formaldehyde reacts with amino acids

Jos J.A.G. Kamps¹, Richard J. Hopkinson ^{1,2*}, Christopher J. Schofield ^{1*} & Timothy D.W. Claridge ^{1*}

Formaldehyde is a biological electrophile produced via processes including enzymatic demethylation. Despite its apparent simplicity, the reactions of formaldehyde with even basic biological components are incompletely defined. Here we report NMR-based studies on the reactions of formaldehyde with common proteinogenic and other nucleophilic amino acids. The results reveal formaldehyde reacts at different rates, forming hydroxymethylated, cyclised, cross-linked, or disproportionated products of varying stabilities. Of the tested common amino acids, cysteine reacts most efficiently, forming a stable thiazolidine. The reaction with lysine is less efficient; low levels of an N^ε-methylated product are observed, raising the possibility of non-enzymatic lysine methylation by formaldehyde. Reactions with formaldehyde are faster than reactions with other tested biological carbonyl compounds, and the adducts are also more stable. The results reveal reactions of formaldehyde with amino acids, and by extension peptides and proteins, have potential roles in healthy and diseased biology, as well as in evolution.

¹Chemistry Research Laboratory, Department of Chemistry, University of Oxford, Mansfield Road, Oxford OX1 3TA, UK. ²Present address: Leicester Institute of Structural and Chemical Biology and School of Chemistry, University of Leicester, Henry Wellcome Building, Lancaster Road, Leicester LE1 7RH, UK.
*email: richard.hopkinson@leicester.ac.uk; christopher.schofield@chem.ox.ac.uk; tim.claridge@chem.ox.ac.uk

Formaldehyde (HCHO), the simplest aldehyde, is present in blood, intercellular tissue, and within cells¹. Exogenous HCHO sources, include pharmaceuticals (including pro-drugs), cosmetics, paper, furniture, cigarette smoke, plywood and foodstuffs^{2,3}. Endogenous cellular HCHO production occurs via enzyme-catalysed reactions involving semicarbazide-sensitive amine oxidases⁴, serine hydroxymethyltransferases⁵, dimethylglycine dehydrogenases⁶, lipid peroxidases⁷, P450 oxidases⁸ and *N*-methyl group demethylases^{9–12}. Recent work also suggests HCHO is involved in folate metabolism¹³. Above threshold levels, HCHO manifests toxicity in humans, with acute HCHO exposure associated with pain, dermatitis¹⁴, nausea, arrhythmia¹⁵, coma and renal failure¹⁶; chronic exposure increases the risk of cancer. Studies with mice have revealed correlations between compromised HCHO metabolism and organ dysfunction, bone marrow failure, leukaemia and liver cancer¹⁷. Impaired HCHO metabolism is lethal in models of Fanconi anaemia, a condition characterised by defects in DNA damage repair¹⁸.

From a chemical perspective, HCHO (in its unhydrated form) is a potent electrophile that can react with biological nucleophiles in proteins and DNA. HCHO reportedly facilitates formation of intra-strand and DNA–protein cross-links *in vitro*^{19,20}. It is probable that related reactions of HCHO in cells are responsible for its toxic/carcinogenic effects, although the precise chemistry of such reactions and their cellular prevalence is poorly defined. Given HCHO is produced endogenously, it is possible that it, or its adducts, play sensing or regulatory roles including in redox metabolism²¹; such roles might include dynamically regulating interactions between biomolecules, and/or forming functionally important adducts to nucleic acids and associated proteins²².

To better understand how HCHO and its derivatives influence cellular functions in health and disease, it is important that its reactions with biomolecules are defined. Von Hippel and McGhee pioneered studies on reactions of nucleosides/nucleotides with HCHO^{19,23}; these, together with more recent work^{18,22}, reveal substantial differences in the rates of reaction of HCHO with nucleobases, and in the stabilities of the HCHO–nucleobase adducts (including when produced by enzymatic oxidation of *N*-methyl groups).

Following pioneering work in the 1920s, numerous studies have been reported on reactions of HCHO with amino acids (AAs) and peptides^{24–38}. However, few detailed systematic studies have been reported on reactions of HCHO with amino acids. Following from recent work investigating some of the products of the reactions of HCHO with arginine, lysine and short peptides^{39,40}, we report systematic nuclear magnetic resonance (NMR) studies on reactions of HCHO with common proteinogenic and other biologically relevant AAs. The results reveal HCHO reacts with different AAs under a range of conditions, giving hydroxymethylated, cyclised, *N*-methylated and *N*-formylated products of very different stabilities. Comparison of the reactions of HCHO with those of other tested biologically relevant aldehydes/ketones implies HCHO reacts fastest and, in general, forms the most stable products. The results highlight the potential of HCHO-derived adducts to be involved in healthy biology, disease and evolution.

Results

Product identification in reactions of AAs and HCHO. Initially, we investigated reactions of HCHO with the following common proteinogenic AAs: alanine (Ala), cysteine (Cys), serine (Ser), threonine (Thr), homocysteine (Hcy), penicillamine (Pen), homoserine (Hse), *allo*-threonine (*allo*-Thr), lysine (Lys), ornithine (Orn), arginine (Arg), histidine (His), tryptophan (Trp), tyrosine (Tyr), asparagine (Asn), glutamine (Gln), *N*^ε-methyllysine (Lys

(Me)), *N*^ε,*N*^ε-dimethyllysine (Lys(Me)₂), 5-hydroxylysine (5hLys) and proline (Pro). Standard reaction conditions comprised AA (4.34 mM) with HCHO (10 equivalents) in D₂O. Reactions were monitored over 48 h by ¹H NMR.

The most reactive proteinogenic AA studied was Cys, which gave thiazolidine **1** as the ‘thermodynamic’ product (Fig. 1, black circle), and hemithioacetal **2** as the apparent ‘kinetic’ product (Fig. 1, white diamond, and Fig. 2 (1)). The observation of hemithioacetal **2** was relatively prolonged under acidic conditions (DCI, 2 equivalents), possibly relating to increased *N*^α-amino group protonation. Cyclisation to give thiazolidine **1** was enhanced at elevated temperatures (308 and 318 K) and on addition of base (NaOD, 1 equivalent). A similar reactivity was observed with homocysteine (Hcy), which gave cyclic thiazinane **3** as the thermodynamic product and hemithioacetal **4** as the kinetic product (Fig. 2 (2))⁴¹. The reaction of the Cys analogue penicillamine (Pen) with HCHO gave a thiazolidine, **5**, and kinetic formation of the hydroxymethylated thiol, **6**. Stable hemiaminal-type adducts or imines were not directly observed within detection limits by reaction of HCHO with the *N*^α-groups of Cys/Hcy/Pen, or with any of the tested AAs; however, hemithioacetals **2/4/6** may not be intermediates in formation of **1/3/5** (see below).

No direct evidence for product formation was accrued in initial reactions between Ser and HCHO under our standard conditions. However, new broad ¹H resonances were identified on addition of base (NaOD, 1 equivalent), implying formation of a metastable product in equilibrium with the Ser starting material. This product was assigned as oxazolidine **7** (Fig. 2 (4)), which is structurally equivalent to thiazolidine **1**⁴². The slower formation rate of **7** from Ser, relative to that of **1** from Cys, correlates with the higher pK_a value of the serinyl hydroxyl group (predicted at 13.4 (ACD/I-Lab, 5.0.0.184, Advanced Chemistry Development, Inc., Toronto, ON, Canada, www.acdlabs.com, (2019)) compared to the cysteinyl thiol group (~8.4 (AAs and acidity values, <http://academics.keene.edu/rblatchly/Chem220/hand/npaa/aawpka.htm>, (accessed 1 May 2019)), suggesting hydroxyl/thiol deprotonation might be (partially) rate limiting³⁹. The apparent rapid interconversion between unreacted Ser and **7**, as indicated by their broadened ¹H resonances, was slowed by cooling (278 K), which enabled structural assignment of **7** (Supplementary Fig. 1).

Investigation into potential stereochemical purity loss under basic conditions, using quantitative ¹H NMR, showed no ²H incorporation (<2%) at the *H*^α position, implying maintenance of stereochemistry (Supplementary Fig. 2).

In contrast to Ser, reaction between Thr and HCHO led to observation of oxazolidine **8** under standard conditions, i.e. without addition of base (Fig. 2 (5))⁴². Conversion to **8** was, however, accelerated by addition of base (NaOD, 1 equivalent) and was concentration dependent. The subtle difference between the Thr secondary alcohol and the Ser primary alcohol thus appears sufficient to favour formation of **8** from Thr, potentially due to a Thorpe–Ingold effect, and/or increase in oxazolidine product stability (**8** compared to **7**). Reaction of HCHO with the Thr epimer *allo*-Thr gave oxazolidine **9** (Fig. 2 (6)) but reaction was less efficient than with Thr. This observation likely reflects a steric clash in the *cis*-substituted oxazolidine ring of **9**.

On addition of HCHO to homoserine (Hse), 1,3-oxazinane **10** was formed (Fig. 2 (7)). Reaction to **10** was observed in the absence of acid or base, suggesting formation of the 1,3-oxazinane is more efficient than formation of 1,3-oxazolidines with Ser, Thr or *allo*-Thr. However, addition of base was required for full conversion to **10**.

The reactions of Asn and Gln with HCHO were considerably slower than those with Cys, Hcy, Pen, Ser, Thr, *allo*-Thr or Hse.

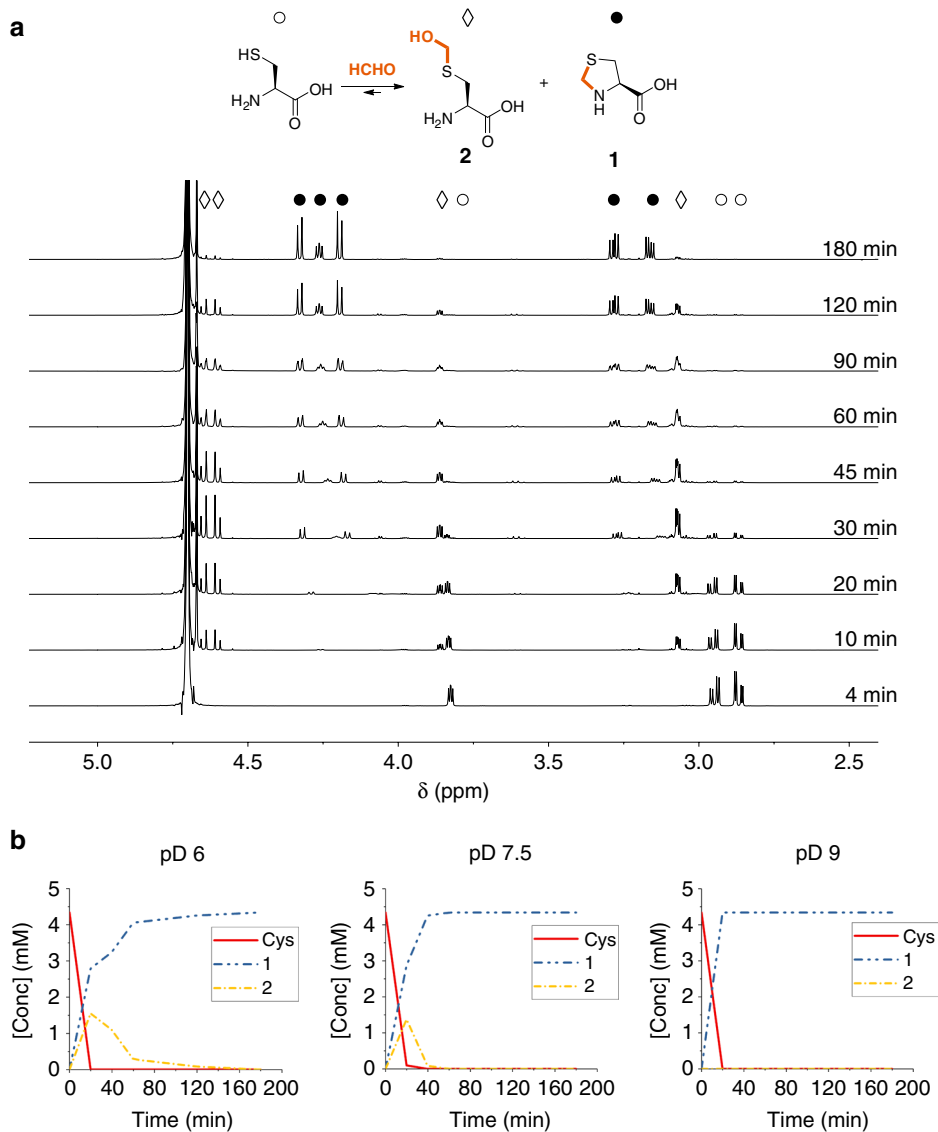


Fig. 1 The reaction between cysteine and formaldehyde (HCHO). HCHO reacts with cysteine to give hemithioacetal and thiazolidine products. **a** Reaction of cysteine (4.34 mM, white circle) with HCHO (2 equiv.) under standard conditions, monitored by ^1H NMR over time, reveals formation of hemithioacetal **2** (white diamond) and thiazolidine **1** (black circle) formation; **b** monitoring the reaction of cysteine (4.34 mM) with HCHO (10 equiv.) under standard conditions, at different pDs: 6, 7.5 and 9

Reaction of Asn and Gln with HCHO over 48 h led to formation of lactams **11** (60%) and **14** (12%), respectively (Fig. 2 (8/9))⁴³. Cyclisation to **11** and **14** was accelerated using more concentrated conditions, by heating, or by addition of base. Conducting the reaction between Asn and HCHO under more concentrated conditions (43 mM Asn; 10 equivalents HCHO) resulted in further hydroxymethylation, forming *N*-hydroxymethylated lactams **12** and **13** (Fig. 2 (8)).

Under standard conditions, His was observed to undergo rapid reaction with HCHO to form the cyclised N^{γ} - and N^{α} -bridged compound **17** (<20 min, Fig. 3 (1)). At higher concentrations, hydroxymethylation of the imidazole ring occurred rapidly, forming **15** and **16**. After 24 h, subsequent formation of the *C*- N^{α} -cyclised molecule spinacine (**19**), a natural product present in many foods^{44,45}, was observed. Formation of **19**, which is likely more stable than **17**, correlated with a decrease in the level of **17**. Formation of **19** was increased with addition of base; however, an

additional hydroxymethylated product **20** was also observed under basic conditions⁴⁵.

Reaction of Trp with HCHO resulted in initial indole ring *N*-hydroxymethylation giving **21**; over time, *C*- N^{α} cyclisation was subsequently observed, affording **22** (Fig. 3 (2)). Addition of acid (DCl, 1 equivalent) suppressed observation of **21** and promoted formation of **22**; addition of NaOD (1 equivalent) promoted formation of **21** over **22**. Further *N*-hydroxymethylation of the indole nitrogen of **22** yielded **23**, which has been reported in reaction of Trp with HCHO under strongly acidic conditions (Fig. 3 (2))⁴⁶.

When Arg was dissolved under standard conditions, a moderately alkaline solution (pH 9) was obtained; reaction with HCHO led to the observation of 1,6,8-triazabicyclodecane **24**, presumably via the transient cyclic intermediate **25** (Fig. 4 (1)), which was initially observed at lower concentrations. Increasing the Arg concentration (to 43.4 mM) resulted in observation of

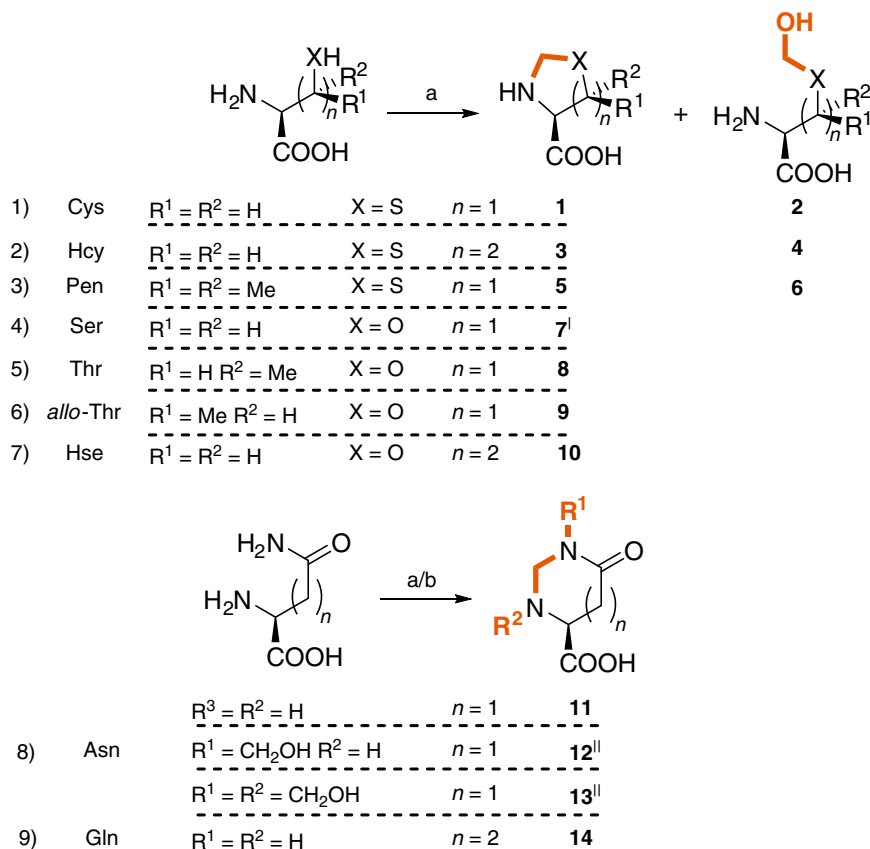


Fig. 2 Products of the reactions of HCHO with thiol/alcohol-containing amino acids. Cysteine, homocysteine, penicillamine, serine, threonine, *allo*-threonine, homoserine, asparagine and glutamine were allowed to react with excess HCHO, in non-buffered D₂O (570 μ L). **a** Amino acid (4.34 mM), HCHO (10 equivalents); **b** amino acid (43.4 mM), HCHO (10 equivalents); (I) required addition of base (NaOD, 1 equivalent); (II) was accelerated by the addition of base (NaOD, 1 equivalent) and at higher concentrations (43.4 mM amino acid, and HCHO)

additional low-level ¹H resonances, which might correspond to acyclic hemiaminals, as previously suggested⁴⁰.

With Orn, reaction was only observed on addition of 1 equivalent of NaOD. At early time-points (5 min), cyclic aminal **26** was the only observed product; **26** is presumably formed via the intermediate aminal **29**, which was observed when exposing Orn to stoichiometric amounts of HCHO (Fig. 4 (2)). Addition of one equivalent of HCHO at high Orn concentrations (43.4 mM) resulted in formation of **29** (31%) and a new species, which was assigned as the dimeric aminal **30** (27%). At a lower Orn concentration (4.34 mM), **29** was observed as the major product (56%). Interestingly, prolonged incubation of Orn (43.4 mM) and HCHO (10 equivalents over 24 h) produced *N*⁶-formyl-*N*¹-methyl and *N*¹-formyl-*N*⁶-methyl adducts **27** and **28** (Fig. 4 (2)); no di-formylated or di-methylated species were observed, suggesting **27** and **28** are formed via intramolecular disproportionation of **26** (likely via an intramolecular hydride shift, as previously proposed for the analogous reaction with simpler alkyldiamines⁴⁷ (Fig. 5a)). After 48 h a small portion (~5%) of **27** and **28** had lost their bridging methylene, affording linear versions of the disproportionated products, i.e., **31** and **32**, respectively. Sufficient levels of **31** and **32** for characterisation were observed after several days. The observations with Orn illustrate the potential for complex and dynamic outcomes in reactions of HCHO with simple AAs.

As with Orn, Lys only reacted with HCHO to form detectable products under basic conditions (NaOD, 1 equivalent), forming low levels of *N*^ε-methylated Lys **33** (<5%, Fig. 4 (3),

Supplementary Figs. 143–146)⁴⁰. No evidence for formation of the

N^α-methylated product was accrued. By contrast with Orn, no cyclic or formylated products were observed with Lys, suggesting the formation of such species is disfavoured when the AA side-chain length increases from 3 to 4 methylenes.

We then explored the reaction of HCHO with 5-hydroxylysine (5hLys, 4.34 mM), which is produced via post-translational modifications in collagen and other proteins (Fig. 4 (4))⁴⁸. No reaction was observed under neutral or acidic conditions with HCHO (10 equivalents); however, new broad resonances appeared with base (1 equivalent of NaOD, Supplementary Fig. 3). The broad nature of the new resonances precluded product characterisation, although it was possible to tentatively assign two resonances to diastereotopic protons of the C6-methylene group of **34** based on chemical shift analysis (δ_{H} 3.02 and 2.41 ppm, respectively, Supplementary Figs. 4 and 5). These two resonances possessed COSY correlations to a single broadened resonance at δ_{H} 3.61 ppm, which was tentatively assigned to the proton attached to C5. HMBC analyses using ¹³C-labelled HCHO revealed correlations between the three resonances and a carbon resonance at δ_{C} 83 ppm, potentially indicative of oxazolidine **34**. Therefore, the available data collectively imply formation of the oxazolidine ring in **34**, as observed with Ser/Thr/*allo*-Thr; the broad nature of the resonances suggests a high degree of reversibility under the tested conditions.

No detectable reaction between Tyr and HCHO was observed. However, previous reports suggest Tyr-residues in peptides can

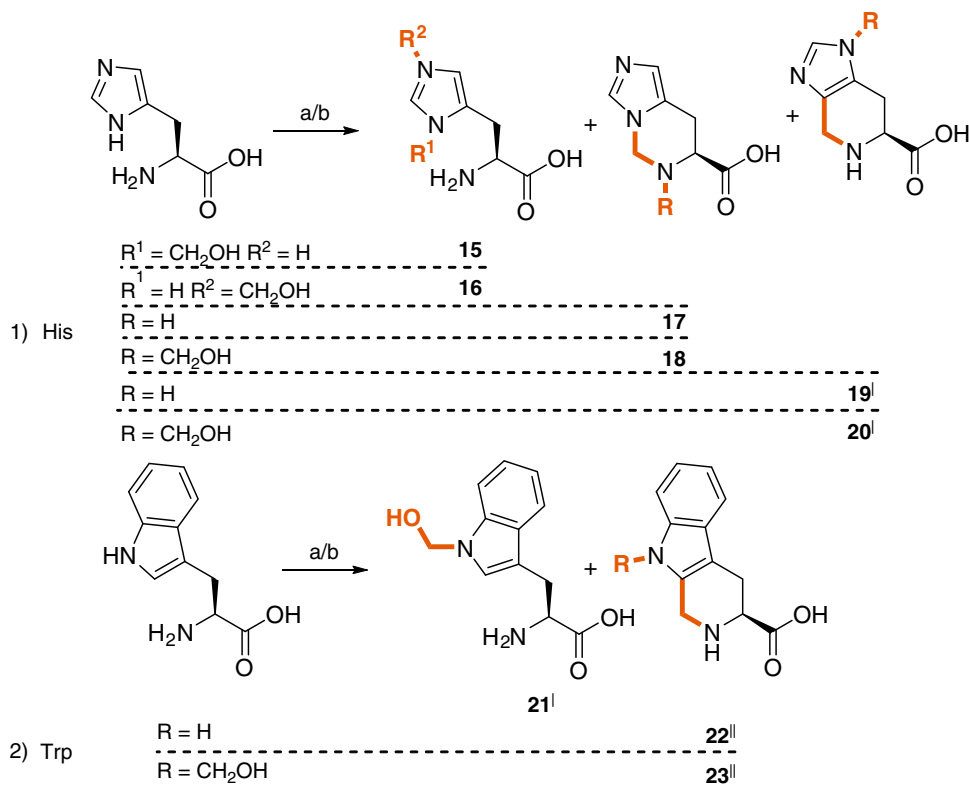


Fig. 3 Products of the reactions of HCHO with aromatic amino acids. Histidine and tryptophan were reacted with HCHO in excess, in non-buffered D₂O (570 μL). **a** Amino acid (4.34 mM), HCHO (10 equivalents); **b** amino acid (43.4 mM), HCHO (10 equivalents); (I) was accelerated under basic conditions (NaOD, 1 equivalent) and at higher concentrations (43.4 mM amino acid); (II) formation of product was promoted under acidic conditions (DCI, 1 equivalent)

react with HCHO and Gly to form bridged Tyr-CH₂-Gly adducts⁴⁹, and HCHO-mediated methylene bridged Tyr-Gly cross-links have been observed in protein crystal structures⁵⁰. No reactions to give stable products were observed between HCHO with either Ala or Pro. These observations imply reactions of the α -amino group to give NMR detectable products require the involvement of additional side-chain groups, at least under our conditions.

pH and concentration dependence studies. Studies then focused on investigating effects of altering the AA concentration and pH. NMR analyses were conducted on samples containing either 4.34 or 43.4 mM AA, pre-adjusted to either pD 6, 7.5 or 9, prior to addition of HCHO (10 equivalents) and monitored up to 48 h. Cys and Hcy reacted with HCHO rapidly at all three pHs and at both concentrations, forming thiazolidines **1** and **3**, respectively, within one hour (Supplementary Table 1 (1/2)). Note that in the absence of HCHO, disulfide formation was observed; thus HCHO has potential to modulate thiol/disulfide equilibria in cells. Ser, Thr, *allo*-Thr and Hse also reacted to form cyclic products (**7–10**), but less efficiently than Cys, reaching equilibrium within 20 min. In all cases cyclisation was greater at the higher tested AA and HCHO concentrations and at pD 9 (Supplementary Table 1 (1–6), 0–86%). When reacting Asn and Gln with HCHO, slow conversion was observed to produce lactams **11** and **14**. Hydroxymethylated lactams **12** and **13** were observed under the more alkaline conditions at both AA concentrations after 48 h (Supplementary Table 1 (7/8)). Formation of **17** by reaction of His and HCHO was preferred at the lower His concentration, and at a lower pH (pD = 6). The C–N cyclised products were more prevalent under basic conditions, with **19** and **20** being observed at

pD 9 after 48 h (Supplementary Table 1 (9)). With Trp, the hydroxymethylated adduct **21** was the major observed product at the lower Trp concentration; with higher Trp concentrations, almost exclusive formation of the cyclised products **22** and **23** was observed after 48 h (Supplementary Table 1 (10)). Arg and Lys remained largely unreactive at both tested AA concentrations and across the tested pH range; the only detectable product with Arg was **24**, which was observed only at pD 9 after 48 h (12% and 23% (4.34 and 43.4 mM of AA, respectively), Supplementary Table 1 (11)). With Lys, trace formation of the *N*^ε-methylated product **33** was observed (Fig. 5c); at pD 9 after 48 h (43.4 mM AA), there was evidence for formation of another, previously undetected adduct (broad ¹H resonance at 6.8 ppm), potentially arising from an imine or hemiaminal (**33b**, Fig. 5c); however, the low concentration of this species precluded full characterisation (Supplementary Table 1 (12), Supplementary Figs. 143–146). No further methylation was observed in reactions of HCHO with *N*^ε-methyllysine (Lys(Me)) or *N*^ε,*N*^ε-dimethyllysine (Lys(Me)₂) over 48 h (43.4 mM AA, 10 equivalents of HCHO, pD 9). Upon exposing Orn to HCHO at pD 9, two *N*-formyl amide conformational isomers of each of **27** and **28** were observed. The formation of **27** and **28** appeared to be accelerated at a higher Orn (43.4 mM) and HCHO concentration and at higher pH (pD 9; Fig. 5b).

To further investigate effects of pH on formation of oxazolidines and oxazinanes, we performed reactions with Ser, Thr, Hse and 5hLys (4.34 mM, Supplementary Figs. 6–11) with HCHO (10 equivalents) at different pHs (25 mM phosphate buffer pD 6.4, 7.0, 8.0, 9.0, 9.8, 11.0 and 12.0). Reactions were equilibrated before analysis. Oxazolidine/oxazinane formation was more prevalent under alkaline conditions; for Ser, formation of oxazolidine **7** was most significant in the sample at pD 12

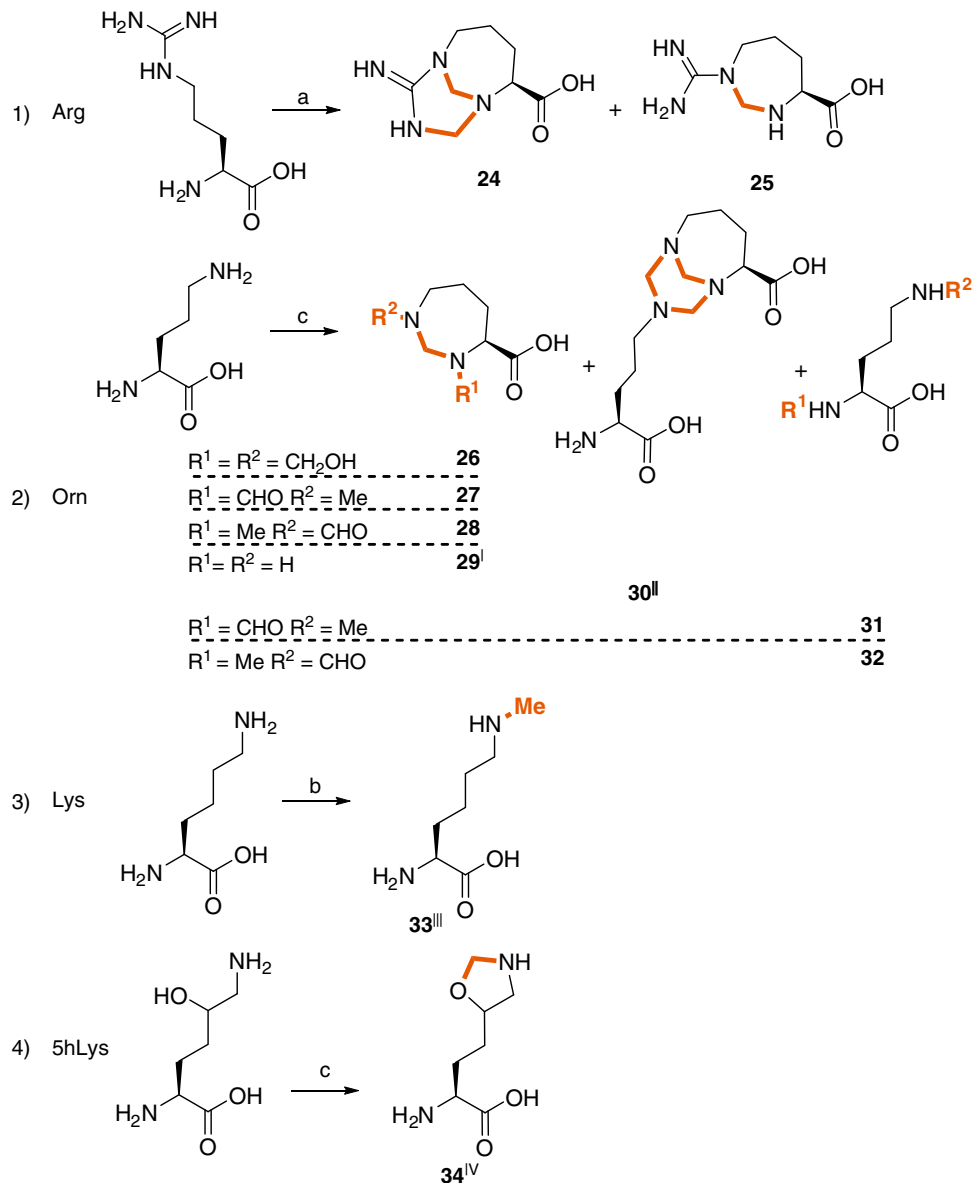


Fig. 4 Products of the reactions of HCHO with sidechain amine-containing amino acids. Arginine, lysine, 5-hydroxylysine and ornithine were allowed to react with HCHO in excess, in non-buffered D₂O (570 μ L). **a** Amino acid (4.34 mM), HCHO (10 equivalents); **b** amino acid (43.4 mM), HCHO (10 equivalents), base (NaOD, 1 equivalent); **c** amino acid (4.34 mM), HCHO (10 equivalents), base (NaOD, 1 equivalent); **d** amino acid (4.34 mM), base (NaOD, 1 equivalent); (I) present at lower concentrations (4.34 mM), and limited HCHO (1 equivalent); (II) present at higher concentrations (43.4 mM), and limited HCHO (1 equivalent); (III) conversion only under alkaline conditions (pD 9) and more product formation at higher concentrations (43.4 mM); (IV) broadened signals prevented full assignment

(>50% conversion), while the Thr-derived oxazolidine **8** was comparably prevalent at pD 8. Similar reactivities were observed with Ser/Thr (2.5 mM) and 20 equivalents of HCHO (Supplementary Fig. 10), suggesting HCHO concentrations are saturating. By contrast, reaction of Hse and HCHO to form oxazinane **10** (>50% conversion) was observed at pD 7, suggesting oxazinane formation is more favourable than oxazolidine formation at neutral pH. Observed reaction of HCHO with 5hLys was most significant at pD 12 (Supplementary Fig. 9); this reactivity is similar to that observed with Ser. Finally, no significant difference in reactivity was observed in samples with Thr or *allo*-Thr (at 2.5 mM, Supplementary Fig. 11).

Various values for *in vivo* HCHO concentrations have been reported^{1,51} likely reflecting difficulties in measuring localised concentrations of a reactive electrophile. To increase the

biological relevance of our results, which we initially carried out with a substantial HCHO excess, we performed reactions with Cys or His at a lower HCHO:AA ratio and at more dilute conditions (AA 25 μ M; HCHO 30 μ M). Despite the dilute concentrations and reduced HCHO excess (1:1.16), Cys still manifested thiazolidine **1** formation; however, no products were observed with His under these conditions (Supplementary Figs. 12 and 13).

Assessing product stabilities. We then investigated the stability of selected AA-HCHO adducts. Firstly, the life-times of the adducts were monitored after removal of residual excess HCHO. This was achieved by either: (i) treatment with the HCHO scavenger 1,3-diketocyclohexane (DCH)⁹, or (ii) lyophilisation followed by re-addition of D₂O. The reaction mixtures of

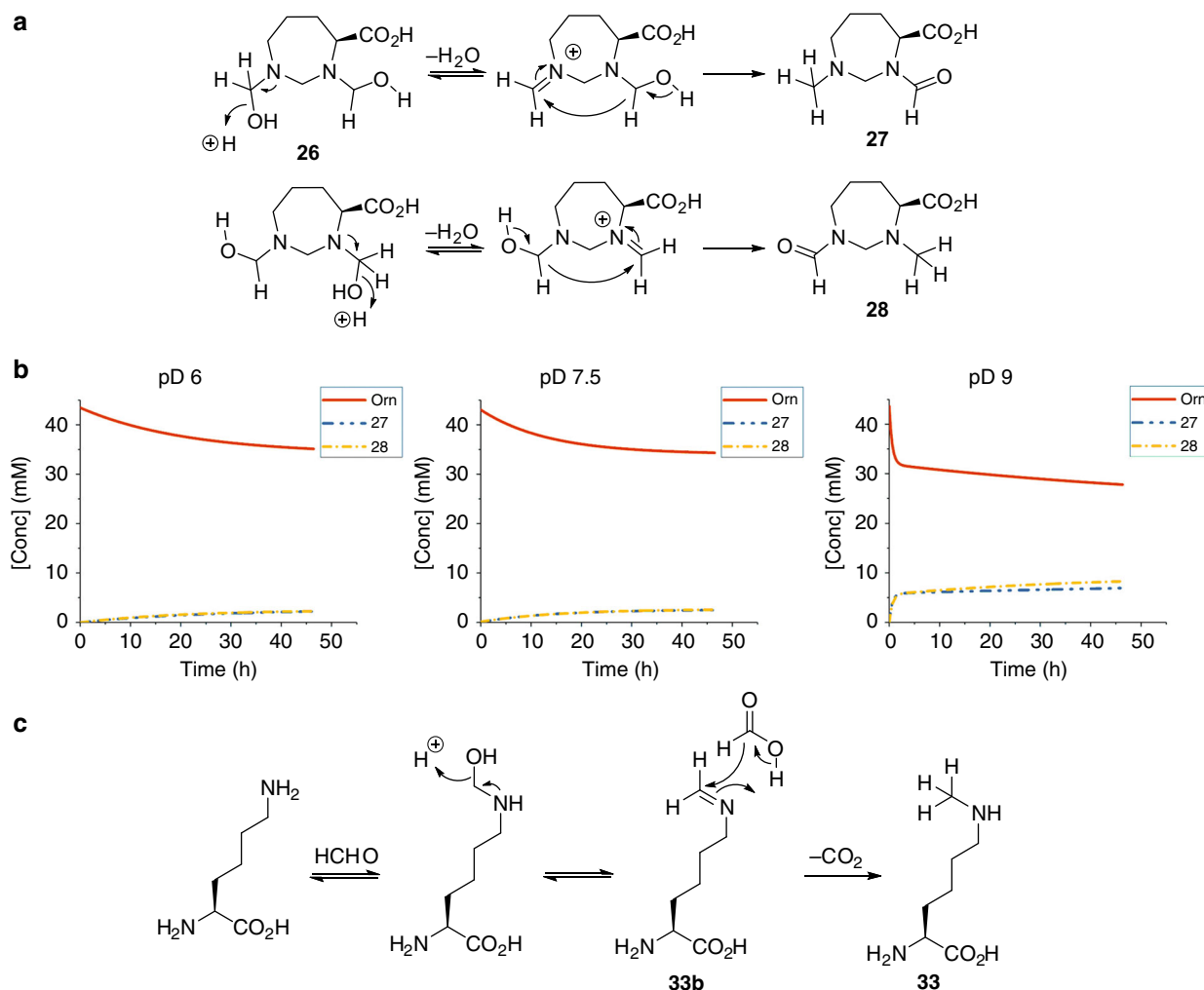


Fig. 5 Possible mechanism and pH profile of the reactions of HCHO with lysine or ornithine. **a** Possible mechanisms for formation of **27** and **28** from **26** (Fig. 4) under basic conditions, (ornithine (4.34 mM), HCHO (10 equiv.) in the presence of base (NaOD, 1 equiv.)). **(b)** 1H NMR time course of the reaction between ornithine (43.4 mM) and HCHO (10 equiv.) at different pHs; concentrations were determined by peak integration and referenced to the original ornithine concentration. **c** possible outline mechanism for formation of **33** via imine **33b**

AA (4.34 mM) and HCHO (10 equivalents) were allowed to reach equilibrium (3–24 h) before HCHO removal. Applying either of the methods to mixtures containing **17** and **18** resulted in rapid re-formation of His, within 1 h, implying fast degradation of the two adducts (Supplementary Figs. 14 and 15). Similarly, fast degradation of oxazolindines **7** and **8** was observed, re-forming Ser and Thr respectively, within 1 and 4 h (Supplementary Figs. 16–18). When mixtures containing the lactams **11** and **12** (from the reaction of Asn and HCHO), and **24** (from Arg and HCHO) were exposed to either of these methods, significantly slower and incomplete formation of the parent AAs was observed (after 72 h), suggesting these lactams (**11**, **12**) are more stable than the cyclised adducts derived from His, Ser and Thr (Supplementary Figs. 19–22). Treating mixtures of Trp-derived **22** and **23** with DCH manifested loss of the N-hydroxymethyl groups; however, the HCHO-derived C,N-linked methylene groups were stable over 48 h (Supplementary Fig. 23). The most stable AA-HCHO adduct of those analysed (Cys adduct **1**, Ser **7**, Thr **8**, Asn **11**, **12** and **13**, Trp **21**, **22** and **23**, Arg **24**) was the Cys-derived thiazolidine **1**, which was stable throughout the analysis time (Supplementary Fig. 24). When ^{13}C -labelled HCHO ($H^{13}CHO$) was added to an aqueous solution of thiazolidine **1**, slow incorporation of the ^{13}C -label into **1** was observed (>48 h)

demonstrating that the product **1** is in slow dynamic equilibrium with Cys (Supplementary Fig. 25). Incorporation of the ^{13}C label was also observed when Trp-derived **21** and **23** were treated with $H^{13}CHO$ over 16 h (Supplementary Fig. 23B); ^{13}C incorporation occurred at the N-hydroxymethyl groups.

Competition between Cys and different AAs for reaction with HCHO. We then investigated competition between AAs for reaction with HCHO, initially focussing on the effects of adding other AAs to reaction mixtures of Cys and HCHO. In competition experiments, where 1 equivalent of HCHO was added to reaction mixtures containing Cys and either Thr, Hse, Asn, His, Trp, Arg, Lys, or Orn (1:1 ratio), exclusive formation of Cys derived thiazolidine **1** was observed within 12 h; with the exception of His, no other products were observed during this time period (Supplementary Figs. 26–34). With His, N,N-cyclised **17** was observed initially, but this product reverted to His alongside the formation of thiazolidine **1** over 5 h.

We then conducted competition experiments with the three thiol-containing AAs, Cys, Hcy and Pen. To a mixture of Cys, Hcy and Pen (1:1:1, at 4.34 mM) was added HCHO (1 equivalent), and the mixture was allowed to react over 6 h (Supplementary Fig. 35). The Pen-derived thiazolidine **5** was

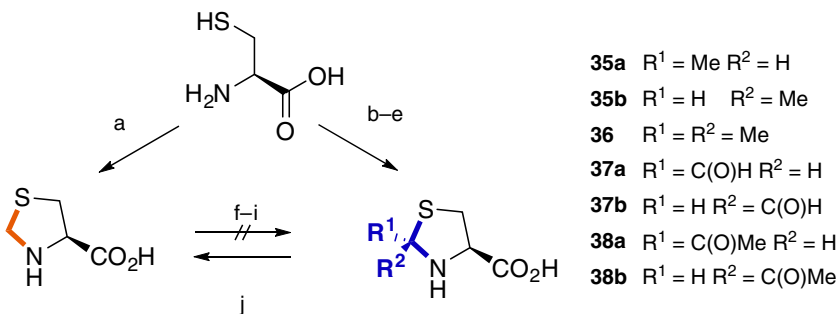


Fig. 6 Products of reactions between cysteine and aldehydes and ketones. Cysteine (4.34 mM) was allowed to react with aldehydes and ketones in non-buffered D₂O (570 μL). **a** HCHO (10 equivalents); **b** acetaldehyde (10 equivalents); **c** acetone (10 equivalents); **d** glyoxal (10 equivalents); **e** methylglyoxal (10 equivalents); **f** acetaldehyde (100 equivalents); **g** acetone (100 equivalents); **h** glyoxal (100 equivalents); **i** methylglyoxal (100 equivalents); **j** HCHO (100 equivalents)

the major product after 6 h (55% of total product), while the Cys-derived thiazolidine **1** (30%) and the Hcy-derived thiazinane **3** (15%) were also observed (Supplementary Fig. 35). These studies suggest that, under the tested conditions, thiazolidine ring formation is favoured over thiazinane ring formation, implying faster formation rates and/or stability. The preferred formation of **5** over **1** is potentially due to the Thorpe–Ingold effect.

Competition experiments with the alcohol-containing AAs, Ser, Thr, *allo*-Thr, and Hse and HCHO (1 equivalent) did not reveal significant formation of adducts. Addition of excess HCHO (100 equivalents) resulted in observable formation of Thr-derived oxazolidine **8** (39%), *allo*-Thr-derived oxazolidine **9** (18%) and Hse-derived oxazinanone **10** (43%, Supplementary Fig. 36). No Ser-derived oxazolidine **7** was observed.

Competition experiments with more complex mixtures of AAs were then conducted. Ser, Thr, *allo*-Thr, Hse, His, Trp, Asn, Orn, Arg and Lys were mixed and reacted with HCHO (1 equivalent) for 24 h. A number of HCHO-derived adducts were observed; the His-derived adduct **17** was the predominant adduct (29%), while compounds **8** (7%), **10** (9%), Asn-derived **11** (10%), Trp-derived **21** (7%) and His-derived **19** (4%) were also observed (some HCHO was unreacted, Supplementary Figs. 37 and 38). On addition of Cys (1 equivalent, 36 h), adducts **8**, **10** and **17** disappeared from the mixture, while formation of **1** was observed. However, adducts **11**, **19** and **21** persisted, implying kinetic stability (Supplementary Figs. 39 and 40, note: these adducts are not observed when Cys was initially present in the mixture, see above).

We then tested whether rates of reaction between Cys and HCHO are affected by the presence of other AAs. Mixtures containing Cys (5 mM), HCHO (1 equivalent) and one of either Ala, Ser, Thr or His (10 equivalents) in phosphate buffer (50 mM, pH 7.5), were monitored over 1 h (Supplementary Fig. 41). With Ala or Ser, only **1** and **2** were observed. With Thr or His, some formation of **8** or **17** was observed, respectively; formation of **8** did not affect the initial formation rate of **1**, while formation of **1** was slowed with addition of His, presumably as a consequence of competing formation of **17**.

The afore-described results imply that thiol containing AAs, are the most reactive with respect to reacting with HCHO. To test whether this apparently special reactivity requires an N^α-amino group, we reacted two tripeptides, glutathione, and L-δ-(2-aminoadipoyl)-L-cysteinyl-D-valine (ACV), under our standard conditions. In agreement with previous studies with glutathione⁵², both these tripeptides formed S-hydroxymethylated adducts at neutral pH (Supplementary Figs. 42 and 43).

Comparison of the reactivity of Cys with HCHO using different electrophiles. Studies were then conducted to investigate the potential reactions of other biologically relevant carbonyl compounds with Cys. Samples were prepared containing Cys and either acetaldehyde (AcH), acetone, glyoxal or methylglyoxal. The reactions produced a set of thiazolidines analogous to the product observed with HCHO (**35a**, **35b**, **36**, **37a**, **37b**, **38a** and **38b**, Fig. 6). Reaction between Cys and acetone gave incomplete formation of **36** (33%), even after addition of base (NaOD, 1 equivalent, 56% of **36**). When reacting with glyoxal to give thiazolidines **37a** and **37b**, formation of the *trans* substituted thiazolidine was moderately preferred (**37a**:**37b**, 1.1:1). Formation of a bithiazolidine was not observed, even with excess Cys (2 equivalents). Methylglyoxal reacted to give thiazolidines **38a** and **38b**, again favouring the *trans*-substituted **38a** (1.7:1).

We then compared the reactivities of the different electrophiles. Cys (4.34 mM) was treated with each of the electrophiles (10 equivalents) and the reaction allowed to react to completion, after which HCHO (100 equivalents) was added. The relative levels of the thiazolidines were then assessed. Displacement of acetone by HCHO from **36** was fastest, resulting in complete formation of thiazolidine **1** within 5 min (Supplementary Figs. 44 and 45). The glyoxal products **37a** and **37b** were displaced to give **1** within an hour, while the methylglyoxal products **38a** and **38b** were present in solution for 6 h after HCHO addition (Supplementary Figs. 46–49). Fragmentation of the acetaldehyde adducts to give thiazolidine **1** was complete in 5 h (Supplementary Figs. 50 and 51). Notably, when the displacement reactions were performed in reverse, i.e., the other electrophiles (100 equivalents) were added to solutions of **1**, no displacement of HCHO from thiazolidine **1** was observed.

Reductive N^α-methylation. Although we did not detect direct evidence for reactions between HCHO and unmodified AA N^α-amino groups under standard conditions, such reactions likely occur (given the nucleophilicity of amino groups) and are potential intermediate steps in the formation of cyclic products. The proposed products of such reactions, i.e., hemiaminals and imines, are presumably unstable/transient, precluding their observation in our NMR analyses. We, therefore, attempted to ‘trap’ N^α-reaction products by well-precedented reductive N-methylation. Cys, Ala and His were individually mixed with HCHO (1 equivalent) and sodium cyanoborohydride (NaCNBH₃, 5 equivalents, Supplementary Figs. 52–54). Notably, with Ala and His, both mono- and di-N^α-methylated products were observed within 1 h, suggesting reactions between HCHO and AA α-amino

groups occur under our aqueous conditions (Supplementary Figs. 53 and 54). With Cys, *N*^α-methylated thiazolidine **1** was observed.

Discussion

We have carried out systematic NMR studies on the reactions of HCHO with common proteinogenic and other AAs, focusing on those with nucleophilic sidechains. The analyses have identified multiple products, revealing the potential for complexity in the reactions of HCHO with even simple biologically relevant small molecules.

Our studies indicate that the *N*^α-amino groups of AAs do not react to give stable (i.e., directly detectable/assignable by NMR) hemiaminal or imine products under our assay conditions. These observations are consistent with previous studies, including work on the mechanism of *N*^ε-methyllysine demethylases, where the proposed intermediate hemiaminals undergo rapid fragmentation⁹. Although we did not directly observe *N*^α-hemiaminal type adducts with any proteinogenic AAs, indirect evidence, including signal broadening, i.e., with Ser, Thr *allo*-Thr, and Hse, suggest these do exist. Indeed, such hemiaminals and the associated imines are likely intermediates in the formation of the cyclised adducts observed by us and others with AAs and peptides^{9,47}. In the case of the reaction of Cys with HCHO, formation of thiazolidine **1** likely proceeds via the unobserved (protonated) *N*^α-linked imine intermediate rather than by the *S*-hydroxymethyl product **2**. Transient imines/hemiaminals are likely intermediates in the efficient reductive *N*-methylation reactions of AAs, as observed by us with Cys, Ala and His, as well as by others with many proteins and peptides using HCHO and NaCNBH₃⁵³.

Some hemiaminals can be observed in aqueous solution, such as the HCHO adduct with Tris buffer and one product derived from the inhibitor/competitive substrate Meldonium after its reaction with γ -butyrobetaine hydroxylase (this hemiaminal is in equilibrium with an oxazolidine)⁵⁴. Thus, since they possess multiple nucleophiles, proteins, nucleic acids, and many other bio-molecules (or complexes thereof) have the potential to sequester HCHO or other reactive carbonyl compounds or to act as reservoirs for those reactive species.

Of the reactive carbonyl compounds we investigated for reaction with Cys, i.e. HCHO, acetone, acetaldehyde, glyoxal and methylglyoxal, HCHO appears to be most reactive, and gives the most stable product (thiazolidine **1**). These observations, which presumably reflect the electrophilicity and small size of HCHO, suggest HCHO has a unique reactivity with biomolecules. A caveat on our work is that in most reactions we used an excess of HCHO and in some cases prolonged reaction times. Different concentration values ($\mu\text{g mL}^{-1}$ range) for *in vivo* HCHO have been reported, likely reflecting difficulties in its analysis. Localised HCHO levels in specific biological environments (e.g., in chromatin) may be higher than generally perceived. Context dependent constraints on concentrations or stereoelectronics may also enhance rates of reaction and/or product stabilities and enzyme (or nucleic acid) catalysis may also be involved. It should also be noted that the high reactivity of HCHO is despite the fact that at neutral pH values HCHO is substantially in its hydrated form, though the carbonyl form is also observed (<0.1%, Supplementary Fig. 55).

Formation of both the R₂N–CH₂–N bridged and RN/S–CH₂–OH hydroxymethyl adducts is reversible, as shown by studies perturbing the equilibrium position by trapping, repeated lyophilisation, or by exchange of the methylenes using H¹³CHO. In some cases, HCHO derived product formation is likely much less reversible or irreversible, e.g., *N*-methylation and

N-formylation (with Lys/Orn), though these products are arguably less biologically relevant, at least in healthy circumstances (see below). The apparent dynamic nature of most identified HCHO-derived adducts renders their identification in cells challenging. Thus, we perceive it is important to define the precise nature of the reaction products of HCHO with isolated biological components with the intention that the results will inform on biological observations.

It is notable that, of all the tested common proteinogenic AAs, Cys reacts most efficiently with HCHO, either as a free AA or in a peptide. This is consistent with the role of the Cys-containing tripeptide glutathione (or other Cys-containing molecules in some organisms) in HCHO detoxification^{13,52}, which proceeds via formation of an *S*-hydroxymethyl adduct. Previous work has found that glutathione can react with HCHO to give cyclic products other than thiazolidines^{52,55}. Although the biological relevance of the glutathione–HCHO-derived and related rings, if any, is unclear, the work with Cys and glutathione reveals the potential for complex reaction outcomes with HCHO and even simple biological molecules. We thus speculate that the reaction of HCHO with biopolymers to form rings may regulate activity (e.g., inhibit proteolysis/alter folding kinetics) or elicit new biochemical functions. The results reveal subtle differences in how the AA side-chains affect the rates of reaction and product stability, e.g. as observed with Thr/*allo*-Thr. As shown by studies with glutathione and another Cys containing tripeptide, we observed that under the tested conditions thiols do not form disulfides when treated with HCHO (Supplementary Figs. 42 and 43). Thus, variations in HCHO and other reactive carbonyl compounds have the potential to directly alter the redox levels in cells via altering thiol/disulfide equilibria. There are several studies showing the potential of histone demethylases to act as hypoxia sensors, a role proposed to be mediated via changes in histone methylation status^{56,57}; our results suggest that HCHO generation by these enzymes should be considered in a regulatory context.

The potential for HCHO to regulate chromatin activity/genetic regulation is of particular current interest given the identification of multiple enzymes catalysing *N*-methyl group demethylation of proteins (principally to date, but not exclusively, histones) and nucleic acids (DNA and RNA) with concomitant HCHO production. Chromatin contains basic elements including nucleobases and histones, which have potential for reaction with HCHO. Indeed, work on the reactions of HCHO with nucleobases has revealed that in some, but not all, cases the *N*-hydroxymethyl adducts are relatively stable (i.e., spectroscopically observable) with substantial half-lives (hours), implying they may accrue over time^{22,58}. Further, HCHO can crosslink the protein and nucleic acid components of chromatin via Mannich-type chemistry^{59,60}.

In the context of chromatin biochemistry, it is interesting that HCHO reacts with basic AAs (Lys, Orn, His and Arg) to give R₂N–CH₂–X-type adducts including rings, and in some cases, *N*-methylated/formylated products. Such non-enzymatic methylation/formylation may be relevant in the context of long term toxic environmental exposure to HCHO, or e.g. by exposure due to treatment of patients with HCHO producing prodrugs^{54,61}. It is also possible that such reactions occur in healthy cells or in the extracellular matrix, due to locally high HCHO levels, e.g. as a consequence of HCHO produced by *N*-methyl demethylase catalysis or during inflammation. As with the other reactions described here, it is also of potential interest with respect to the early stages of the evolution of current biochemistry.

In the case of some HCHO reactions with AAs with *N*-containing side-chains (e.g., Arg, Orn, His and Trp), formation of ring systems was observed; over time, His and Trp underwent

cyclisation forming a C–N bond via Pictet–Spengler type reaction (Fig. 3). The complex rings systems formed with Orn and Arg are particularly notable (Fig. 4).

It is also important to note that, as well as modulating interactions between biopolymers by formation of covalent links, both stable or transient HCHO adducts have the potential to modify biomolecular properties/functions (e.g., protein/nucleic acid folding) by altering non-covalent interactions, altering solvation patterns²², or by modulating HCHO availability, as evidenced by our competition experiments.

Our overall results reveal that even the reactions of HCHO with AAs can have complex and dynamic outcomes. Combined with other results, they reveal the potential for significant biochemical complexity in the reactions of HCHO and potentially other electrophiles with the nucleophilic components of chromatin and other biological systems. The reactions of HCHO with histones tails, which contain a high density of basic/nucleophilic residues, are the subject of ongoing investigations. Whether or not HCHO and its reactions on chromatin and other cellular systems are involved in the regulation of normal healthy biology is presently an open question. However, it is clear that, at least above threshold levels, exogenous HCHO can promote disease, e.g. cancer⁶². HCHO is also used to modulate pain response in animal models, likely at least in part via reaction with cysteine residues in transient receptor potential sensory protein. HCHO is used as a preservative^{63,64}, and its derivatives are also present in food (e.g., spinacine), prodrugs^{54,61}, drug metabolites (e.g., of Meldonium⁵⁴) and in natural products (e.g., tubulysin)^{65,66}. It is likely some of the reactions studied here are relevant to these roles of HCHO.

Methods

General methods. Alanine monohydrate, arginine, asparagine, 1,3-cyclohexadione, cysteine, dimedone, 5-hydroxylysine hydrochloride salt, homocysteine, homoserine, lysine hydrochloride salt, *N*^ε-methyllysine hydrochloride salt, *N*^ε,*N*^ε-dimethyllysine hydrochloride salt, ornithine hydrochloride salt, proline, serine, threonine, *allo*-threonine, tryptophan and tyrosine, were from commercial suppliers and used without purification. Samples were prepared in D₂O (purity 99.9 atom% D), unless stated otherwise. Starting materials and products formed during the reactions were monitored by NMR and characterised using a combination of 1D and 2D techniques. Samples were recorded at 298 K, unless stated otherwise, in a 5 mm tube, using a Bruker AVII 500 equipped with a TXI H/F/C probe, Bruker AVIII HD 500 equipped with a BBFO probe, Bruker AVII 500 equipped with a CPDUL He CryoProbe, Bruker AVIII HD 600 equipped with a BB-F/H N₂ CryoProbe or a Bruker AVIII 700 equipped with a TCI H/C/N He CryoProbe, at their respective frequencies.

Standard conditions reaction setup. AAs (0.0247 mmol) were weighed and dissolved in D₂O (1 mL)–25 mM AA stock solution. HCHO (30.0 mg, 1.0 mmol) was suspended with D₂O (3 mL) and heated with a hairdryer until a clear colourless solution was obtained—the 335 mM HCHO stock solution. In the cases of acetaldehyde, acetone, glyoxal or methylglyoxal an equimolar solution was created by addition of an appropriate volume of one of the compounds to D₂O (3 mL). A typical sample consisted of the AA stock solution (100 μL, 2.47 μmol, 1 equivalent) and D₂O (400 μL); the solution was thoroughly mixed and the NMR spectrum was recorded. Optionally, DCl (1 M) or NaOD (1 M) were added, mixed and recorded. Subsequently HCHO (7–70 μL, 2.47–24.7 μmol, 1–10 equivalents) was added and the progress of the reaction was monitored over time. Variable temperature (VT) studies required the samples to be adjusted to the appropriate temperature before the addition of electrophile. Competition experiments were conducted using AA1 stock solution (100 μL), AA2 stock solution (100 μL) and D₂O (300 μL), with a limited amount of HCHO (1 equivalent).

Adjusted pH reactions. The appropriate stock solution was diluted to a minimum volume of 4 mL and the pH was adjusted to: 5.60, 7.10 or 8.60, corresponding to pD 6.00, 7.50 or 9.00, using NaOD and DCl (0.1 M and 0.01 M). HCHO stock solution was prepared by suspending paraformaldehyde (30.0 mg, 1 mmol) in D₂O and subsequently adjusting the pH with NaOD and DCl (0.1 M and 0.01 M) to the desired pH (total volume eventually 3 mL). AA stock solution (400 μL, 2.47 μmol) was mixed with D₂O (100 μL); and NMR spectra were recorded. Subsequently, HCHO (7–70 μL, 2.46–24.6 μmol, 1–10 equivalents) was added and reaction

progress was monitored over time (maximum 48 h). See Supplementary Fig. 55 for analysis of the stock solution

pH dependent equilibria experiments. To a buffered solution (144 μL, 25 mM phosphate buffer, at pH = 5.9, 6.6, 7.6, 8.6, 9.4, 10.6 and 11.6), an AA (16 μL, 50 mM stock solution; in the case of 2 AAs: 2 × 8 μL was used) was added. Subsequently, HCHO (2.38 μL, 3.35 M, 10 equivalents), was added. The mixture was allowed to equilibrate for at least 20 min before recording the NMR spectrum.

Stability assays using 1,3-cyclohexadione. Upon completion of a standard reaction, as monitored by ¹H NMR, a solution of 1,3-cyclohexadione (22.4 mg, 0.2 mmol, in 180 μL, 1.1 M; 85 μL, 40 equivalents) was added to the reaction mixture. Formation of the dimerised diketone and parent AA formation were both followed by NMR over time.

Lyophilisation to test stability of HCHO adducts. Upon completion of a standard reaction, as monitored by ¹H NMR, the solution was frozen in liquid N₂ and lyophilised overnight. The mixture was then redissolved in D₂O and allowed to equilibrate for 1 h and analysed by NMR. The sample was then put through a cycle of lyophilisation and redissolving several times (5–8).

Using ¹³C labelled HCHO to test for exchange of HCHO adducts. A reaction mixture with AA was stirred under standard conditions using unlabelled HCHO (5 equivalents), until the species of interest was observed. Subsequently, labelled H¹³CHO (5 equivalents) was added to the solution and progress was monitored over time.

Competition assays. A stock solutions of one AA (25 mM) was prepared and added (100 μL) to D₂O (300 μL). Cysteine (100 μL, or another AA) was then added, to give an equimolar solution (50 mM, 1:1). This was followed by the addition of HCHO (1 equivalent, 3.35 M stock solution). Reactions were monitored by ¹H NMR for 16 h until all HCHO was consumed.

For competition assays with more than four AAs, higher stock solution concentrations (200 mM) were used.

Spectator AA experiments. A solution of cysteine (5 mM) and a ‘spectator’ AA (25 mM), in phosphate buffer (50 mM, pH 7.10, in >95% D₂O) was made and HCHO (1 equivalent with respect to cysteine, 3.35 M stock solution) was added. The reaction was followed by ¹H NMR.

Reductive amination. A mixture of AA (5 mM) and NaBH₃CN (5 equivalents) in D₂O was prepared. Subsequently, HCHO (1.1 equivalents, 3.35 M stock solution) was added. The reaction was followed and upon completion analyzed by ¹H NMR.

Product characterisation. Full spectra of the following compounds are provided in Supplementary Note 1 and Supplementary Figs. 57–146.

Cysteine products. (S)-4-Carboxy-1,3-thiazolidine 1: The title compound was observed when cysteine (4.34 mM) was exposed to HCHO (10 equiv.) and left to react for 1 h. ¹H NMR (700 MHz, D₂O) δ: 4.33 (d, 10.0 Hz, 1H, H₂), 4.29 (dd, *J* = 7.5, 6.0 Hz, 1H, H₄), 4.21 (d, *J* = 10.0 Hz, 2H, H₂), 3.29 (dd, 1H, 12.0, 7.5 Hz), 3.18 (dd, 1H, 12.0, 7.5 Hz). ¹³C NMR (176 MHz, D₂O) δ: 172.0, 81.6, 63.9, 48.6, 32.9. Electrospray ionisation mass spectrometry (ESI–MS) calculated for C₄H₈O₂NS [M + 1]⁺ 134.02703, found 134.02704. These results are in agreement with previously reported data⁶⁷.

(S)-Hydroxymethyl-L-cysteine 2: The title compound was observed after 30 min, when adding acid (DCl, 2 equiv.) to Cys (4.34 mM) followed by HCHO (2 equiv.) addition. ¹H NMR (700 MHz, D₂O) δ: 4.62 (q, *J* = 12.0 Hz, 2H, H₈), 4.22 (dd, *J* = 7.0, 4.0 Hz, 1H, H₂), 3.15 (dd, *J* = 15.5, 4.0 Hz, 2H, H₆). ¹³C NMR (176 MHz, D₂O) δ: 170.4, 81.6, 65.1, 52.7, 31.0.

Homocysteine products. (S)-4-Carboxy-1,3-thiazanane 3: The title compound was observed when homocysteine (4.34 mM) was exposed to HCHO (10 equiv.) and left to react for 1 h. ¹H NMR (700 MHz, D₂O) δ: 4.16 (d, *J* = 13.0 Hz, 1H, H₂), 4.10 (d, *J* = 13.0 Hz, 1H, H₂), 3.52 (dd, *J* = 12.0, 3.0 Hz, 1H, H₄), 2.85 (dt, *J* = 14.0, 3.0 Hz, 1H, H₆), 2.69 (dt, *J* = 14.0, 3.0 Hz, 1H, H₆), 2.39 (dq, *J* = 15.0, 3.5 Hz, 1H, H₂), 1.87 (dq, *J* = 15.0, 3.5 Hz, 1H, H₂). ¹³C NMR (176 MHz, D₂O) δ: 170.0, 59.4, 45.2, 28.1, 26.0. These data are in agreement with those reported⁶⁸.

(S)-Hydroxymethyl-L-homocysteine 4: The title compound was observed when adding acid (DCl, 2 equiv.) to Hcy (4.34 mM) followed by HCHO (2 equiv.) addition, after 30 min. ¹H NMR (500 MHz, D₂O) δ: 4.67 (s, 2H, H₉), 4.15 (t, *J* = 6.5 Hz, 1H, H₂), 2.75 (t, *J* = 7.5 Hz, 2H, H₇), 2.28 – 2.21 (dt, *J* = 14.5, 7.5 Hz, 1H, H₃), 2.14 (dt, *J* = 14.5, 7.5 Hz, 1H, H₃). ¹³C NMR (126 MHz, D₂O) δ: 171.6, 64.6, 51.6, 30.0, 25.7.

D₂O) δ : 181.1, 157.0, 67.9, 64.8, 63.9, 50.0, 28.8, 28.2. ESI-MS calculated for C₈H₁₅O₂N₄ [M + 1]⁺ 199.11895, found 199.11892.

Ornithine products. (S)-1,3-Bis(hydroxymethyl)-1,3-diazepane-4-carboxylic acid **26**: The title compound was observed when ornithine (HCl salt) (43.4 mM) was exposed to HCHO (10 equiv.) and base (NaOD, 2 equiv.) after 3 min. ¹H NMR (700 MHz, D₂O) δ : 4.81 (d, *J* = 11.0 Hz, 1H, H₁₁), 4.75 (d, *J* = 4.7 Hz, 1H, H₁₂), 4.49 (d, *J* = 11.0 Hz, 1H, H_{11'}), 4.36 (d, *J* = 11.0 Hz, 1H, H_{12'}), 4.14 (d, *J* = 14.5 Hz, 1H, H₆), 3.87 (d, *J* = 14.5, 1H, H₆), 3.37 (dd, *J* = 12.0, 4.5 Hz, 1H, H₁), 2.97 (dt, *J* = 14.5, 2.0 Hz, 1H, H₄), 2.65 (t, *J* = 14.0 Hz, 1H, H₄), 2.21–2.13 (m, 1H, H₂), 1.98–2.05 (m, 1H, H₃), 1.84–1.78 (m, 1H, H₃), 1.73–1.65 (m, 1H, H₂). ¹³C NMR (176 MHz, D₂O) δ : 183.0, 86.2, 84.9, 69.1, 62.7, 55.1, 32.0, 29.2.

(S)-3-Formyl-1-methyl-1,3-diazepane-4-carboxylic acid **27**: The title compound was observed when ornithine (HCl salt) (43.4 mM) was exposed to HCHO (10 equiv.) and base (NaOD, 2 equiv.) after 2 h. ¹H NMR (600 MHz, D₂O) δ (two isomers, C and D): 8.30 (s, 1H, H_{C11}), 8.15 (d, *J* = 2.0 Hz, 1H, H_{D11}), 5.14 (d, *J* = 13.0 Hz, 1H, H_{D2}), 4.77–4.73 (m, 10H, H_{C2}), 4.62 (d, *J* = 14.0 Hz, 1H, H_{C2}), 4.34–4.26 (m, 2H, H_{C4/D4/D2}), 3.42–3.32 (m, 2H, H_{A7/D7}), 3.21 (d, *J* = 13.5 Hz, 1H, H_{C7}), 2.97 (t, *J* = 13.0 Hz, 1H, H_{D7}), 2.86 (t, *J* = 13.0 Hz, 1H, H_{C7}), 2.80 (s, 4H, H_{C4/D4}), 2.61 (s, 3H, H_{C14/D5}), 2.48–2.42 (m, 1H, H_{C5}), 2.08–2.02 (m, 1H, H_{A5'/D6}), 2.00–1.94 (m, 1H, H_{C6}), 1.93–1.76 (m, 4H, H_{B6/C5}), 1.76–1.65 (m, 2H, H_{C6}). ¹³C NMR (151 MHz, D₂O) δ : 177.9 (D8), 177.9 (C8), 167.2 (C11), 166.7 (D11), 65.9 (C2), 61.2 (D4), 60.1 (D2), 59.5 (C4), 58.3 (D7), 57.1 (C7), 40.3 (D14), 39.0 (C14), 29.4 (D5), 28.7 (C5), 23.4 (C6), 22.4 (D6).

(S)-1-Formyl-3-methyl-1,3-diazepane-4-carboxylic acid **28**: The title compound was observed when ornithine (HCl salt) (43.4 mM) was exposed to HCHO (10 equiv.) and base (NaOD, 2 equiv.) after 2 h. ¹H NMR (600 MHz, D₂O) δ (two isomers, A and B): 8.21 (s, 1H, H_{A11}), 8.14 (d, *J* = 2.0 Hz, 1H, H_{B11}), 4.96 (dd, *J* = 14.0, 2.0 Hz, 1H, H_{A7}), 4.86–4.81 (m, 2H, H_{B2}), 4.70 (s, 24H, H_{B2}), 4.54 (d, *J* = 14.0 Hz, 1H, H_{A7}), 3.72–3.66 (m, 1H, H_{B4}), 3.63–3.59 (m, 1H, H_{A4}), 3.56–3.48 (m, 3H, H_{A7/B7}), 3.42–3.32 (m, 2H, H_{A7/D7}), 2.64 (s, 3H, H_{A14}), 2.22–2.17 (m, 2H, H_{B5}), 2.17–2.10 (m, 1H, H_{A5}), 2.08–2.02 (m, 1H, H_{A5'/D6}), 1.93–1.76 (m, 4H, H_{B6/C5}). ¹³C NMR (151 MHz, D₂O) δ : 175.5 (A8), 173.9 (B8), 167.6 (B11), 166.8 (A11), 85.7 (B2), 70.5 (B4), 69.1 (A4), 66.1 (A7), 60.8 (B2), 48.6 (B7), 44.1 (A7), 41.3 (B14), 38.9 (A14), 26.1 (B6), 25.9 (A5), 25.3 (B5), 23.3 (A6).

(S)-1,3-Diazepane-4-carboxylic acid **29**: The title compound was observed when ornithine (HCl salt) (43.4 mM) was exposed to HCHO (1 equiv.) and base (NaOD, 2 equiv.) after 10 min. ¹H NMR (700 MHz, D₂O) δ : 3.71 (d, *J* = 13.5 Hz, 1H, H₂), 3.40 (d, *J* = 13.5 Hz, 1H, H₂), 3.22 (dd, *J* = 9.0, 4.0 Hz, 1H, H₄), 2.70–2.59 (m, 2H, H₇), 1.66–1.51 (m, 2H, H₅), 1.50–1.40 (m, 2H, H₆). ¹³C NMR (176 MHz, D₂O) δ : 182.2, 61.2, 60.0, 46.0, 32.5, 27.9.

(1R,2S,6R)-8-(S)-4-Amino-4-carboxybutyl)-1,6,8-triazabicyclo[4.3.1]decane-2-carboxylic acid **30**: The title compound was observed when ornithine (HCl salt) (43.4 mM) was exposed to HCHO (1 equiv.) and base (NaOD, 2 equiv.) after 20 min. ¹H NMR (700 MHz, D₂O) δ : 3.85 (d, *J* = 14.3, 1H, H₂), 3.72 (d, *J* = 11.5, 1H, H₁₁), 3.54 (d, *J* = 11.5 Hz, 1H, H₁₂), 3.29 (dd, *J* = 12.0, 4.1 Hz, 1H, H₄), 3.19 (d, *J* = 11.5 Hz, 1H, H₁₁), 3.10 (m, *J* = 11.5 Hz, 2H, H_{12'} and H₁₇), 2.91–2.85 (m, 1H, H₇), 2.65–2.59 (m, 1H, H₇), 2.20–2.12 (m, 2H, H₆), 1.69–1.62 (m, 2H, H₁₆), 1.55–1.46 (m, 2H, H₁₄), 1.35–1.28 (m, 2H, H₁₅). ¹³C NMR (176 MHz, D₂O) δ : 183.6, 183.5, 76.1, 73.8, 69.1, 63.5, 55.0, 51.8, 32.8, 31.7, 28.5, 21.6.

(S)-2-Formamido-5-(methylamino)pentanoic acid **31**: The title compound was observed when ornithine (HCl salt) (43.4 mM) was exposed to HCHO (10 equiv.) and base (NaOD, 2 equiv.) after several days and more build up over a period of a week. ¹H NMR (700 MHz, D₂O) δ : 8.04 (s, 1H, H₁₀), 4.32 (dd, *J* = 8.0, 5.0, 1H, H₂), 3.01–2.96 (m, 2H, H₆), 2.62 (s, 3H, H₁₃), 1.89–1.84 (m, 1H, H₄), 1.75–1.68 (m, 1H, H₄), 1.70–1.63 (m, 2H, H₅). ¹³C NMR (176 MHz, D₂O) δ : 176.2, 163.9, 51.90, 48.4, 32.6, 28.1, 21.8.

(S)-5-Formamido-2-(methylamino)pentanoic acid **32**: The title compound was observed when ornithine (HCl salt) (43.4 mM) was exposed to HCHO (10 equiv.) and base (NaOD, 2 equiv.) after several days and more build up over a period of a week. ¹H NMR (700 MHz, D₂O) δ : 7.96 (s, 1H), 3.56–3.53 (m, 1H, H₁), 3.20–3.17 (m, 2H, H₂), 2.62 (s, 3H, H₁₃), 1.85–1.75 (m, 2H, H₃), 1.57–1.44 (m, 2H, H₄). ¹³C NMR (176 MHz, D₂O) δ : 173.2, 164.3, 63.0, 37.3, 31.6, 24.0.

Lysine product. N^ε-Methyl-L-lysine **33**: The title compound was observed when lysine (HCl salt) (43.4 mM) was exposed to HCHO (10 equiv.) at pH 9 after 48 h. ¹H NMR (700 MHz, D₂O) δ : 3.61 (t, *J* = 6.1 Hz, 1H, H₂), 2.92 (t, *J* = 10.0 Hz, 2H, H₉), 2.56 (s, 3H, H₁₁), 1.80–1.71 (m, 2H, H₃), 1.62–1.55 (m, 2H, H₃), 1.40–1.24 (m, 2H, H₇). ¹³C NMR (176 MHz, D₂O) δ : 174.5, 54.4, 48.5, 32.6, 29.8, 26.3, 21.3.

These results are in agreement with previously reported data⁴⁰. Full spectra of the following compounds are provided in Supplementary Note 2 and Supplementary Figs. 148–160.

Cysteine and acetaldehyde products. (2S,4R)-2-Methylthiazolidine-4-carboxylic acid **35a**: The title compound was observed when cysteine (43.4 mM) was exposed to acetaldehyde (10 equiv.) and base (NaOD, 1 equiv.) after 1 h. ¹H NMR (700 MHz, D₂O) δ : 4.82 (q, *J* = 6.5 Hz, 1H), 4.37 (t, *J* = 7.1 Hz, 1H), 3.40 (dd, *J* = 12.1, 7.6 Hz, 1H), 3.19 (dd, *J* = 12.1, 6.6 Hz, 1H), 1.53 (d, *J* = 6.5 Hz, 4H). ¹³C NMR (176 MHz, D₂O) δ : 171.8, 63.49, 60.5, 32.6, 18.3.

(2R,4R)-2-Methylthiazolidine-4-carboxylic acid **35b**: The title compound was observed when cysteine (43.4 mM) was exposed to acetaldehyde (10 equiv.) and base (NaOD, 1 equiv.) after 1 h. ¹H NMR (700 MHz, D₂O) δ : 4.74 (q, *J* = 6.4 Hz, 1H), 4.29 (dd, *J* = 7.7, 6.7 Hz, 1H), 3.34 (dd, *J* = 12.0, 7.8 Hz, 1H), 3.24 (dd, *J* = 12.0, 6.7 Hz, 1H), 1.55 (d, *J* = 6.5 Hz, 3H). ¹³C NMR (176 MHz, D₂O) δ : 171.8, 64.3, 61.0, 32.8, 17.3.

Cysteine and acetone product. (R)-2,2-dimethylthiazolidine-4-carboxylic acid **36**: The title compound was observed when cysteine (43.4 mM) was exposed to acetone (10 equiv.) and base (NaOD, 1 equiv.) after 3 h. ¹H NMR (600 MHz, D₂O) δ : 4.52 (t, *J* = 7.5 Hz, 1H, H₅), 3.57 (dd, *J* = 12.5, 8.0 Hz, 1H, H₁), 3.41 (dd, *J* = 12.5, 8.0 Hz, 1H, H₁), 1.76 (d, *J* = 11.0 Hz, 6H, H₉ and H₁₀). ¹³C NMR (151 MHz, D₂O) δ : 172.3, 73.2, 64.3, 32.9, 27.5, 26.9.

Cysteine and glyoxal products. (2S,4R)-2-(2-Oxoacetyl)thiazolidine-4-carboxylic acid **37a**: The title compound was observed when cysteine (43.4 mM) was exposed to glyoxal (10 equiv.) after 1 h. ¹H NMR (700 MHz, D₂O) δ : 5.15 (d, *J* = 4.8 Hz, 1H, H₁₁), 4.46 (dd, *J* = 7.0, 4.9 Hz, 1H, H₃), 3.30 (dd, *J* = 11.3, 6.9 Hz, 1H, H₄), 3.21 (dd, *J* = 12.1, 5.0 Hz, 1H, H₄). ¹³C NMR (176 MHz, D₂O) δ : 171.3, 88.6, 66.9, 64.6, 32.2.

(2R,4R)-2-(2-Oxoacetyl)thiazolidine-4-carboxylic acid **37b**: The title compound was observed when cysteine (43.4 mM) was exposed to HCHO (10 equiv.) after 1 h. ¹H NMR (700 MHz, D₂O) δ : 5.19 (d, *J* = 6.2 Hz, 1H, H₁₁), 4.39 (t, *J* = 7.2 Hz, 1H, H₃), 3.33 (d, *J* = 7.2 Hz, 1H, H₄), 3.25–3.23 (m, 1H, H₄). ¹³C NMR (176 MHz, D₂O) δ : 171.1, 88.3, 67.6, 64.9, 32.3.

Cysteine and methylglyoxal products. (2S,4R)-2-Acetylthiazolidine-4-carboxylic acid **38a**: The title compound was observed when cysteine (43.4 mM) was exposed to methylglyoxal (10 equiv.) after 1 h. ¹H NMR (600 MHz, D₂O) δ : 5.65 (s, 1H, H₁), 4.45 (dd, *J* = 9.8, 6.6 Hz, 1H, H₃), 3.44–3.41 (m, 1H, H₄), 2.99 (dd, *J* = 12.5, 9.7 Hz, 1H, H₄), 2.34 (s, 3H, H₁₁). ¹³C NMR (151 MHz, D₂O) δ : 200.4, 170.8, 67.8, 66.8, 32.9, 25.1.

(2R,4R)-2-Acetylthiazolidine-4-carboxylic acid **38b**: The title compound was observed when cysteine (43.4 mM) was exposed to methylglyoxal (10 equiv.) after 1 h. ¹H NMR (600 MHz, D₂O) δ : 5.72 (s, 1H, H₁), 4.55 (dd, *J* = 7.6, 3.3 Hz, 1H, H₃), 3.40–3.37 (m, 1H, H₄), 3.18 (dd, *J* = 12.3, 7.2 Hz, 1H, H₄), 2.31 (s, 3H, H₁₁). ¹³C NMR (151 MHz, D₂O) δ : 200.7, 171.6, 67.7, 64.8, 32.9, 24.9.

Data availability

All data generated and analysed during this study are included in this published article and its Supplementary information file.

Received: 7 June 2019; Accepted: 23 September 2019;

Published online: 07 November 2019

References

- Szarvas, T. et al. Determination of endogenous formaldehyde level in human blood and urine by dimedone-14C radiometric method. *J. Radioanal. Nucl. Chem.* **106**, 357–367 (1986).
- Zhang, L., Steinmaus, C., Eastmond, D. A., Xin, X. K. & Smith, M. T. Formaldehyde exposure and leukemia: a new meta-analysis and potential mechanisms. *Mutat. Res.* **681**, 150–168 (2009).
- Gillman, I. G., Kistler, K. A., Stewart, E. W. & Paolantonio, A. R. Effect of variable power levels on the yield of total aerosol mass and formation of aldehydes in e-cigarette aerosols. *Regul. Toxicol. Pharmacol.* **75**, 58–65 (2016).
- Boor, P. J., Trent, M. B., Lyles, G. A., Tao, M. & Ansari, G. A. S. Methylamine metabolism to formaldehyde by vascular semicarbazide-sensitive amine oxidase. *Toxicology* **73**, 251–258 (1992).
- Schirch, V. & Szebenyi, D. M. E. Serine hydroxymethyltransferase revisited. *Curr. Opin. Chem. Biol.* **9**, 482–487 (2005).
- Steenkamp, D. J. & Husain, M. The effect of tetrahydrofolate on the reduction of electron transfer flavoprotein by sarcosine and dimethylglycine dehydrogenases. *Biochem. J.* **203**, 707–715 (1982).
- Orrenius, S., Dallner, G. & Ernster, L. Inhibition of the TPNH-linked lipid peroxidation of liver microsomes by drugs undergoing oxidative demethylation. *Biochem. Biophys. Res. Commun.* **14**, 329–334 (1964).
- Teschke, R. & Lieber, S. Hepatic microsomal system: and characterization. *Arch. Biochem. Biophys.* **415**, 404–415 (1974).
- Hopkinson, R. J., Hamed, R. B., Rose, N. R., Claridge, T. D. W. & Schofield, C. J. Monitoring the activity of 2-oxoglutarate dependent histone demethylases by NMR spectroscopy: direct observation of formaldehyde. *ChemBioChem* **11**, 506–510 (2010).
- Walport, L. J., Hopkinson, R. J. & Schofield, C. J. Mechanisms of human histone and nucleic acid demethylases. *Curr. Opin. Chem. Biol.* **16**, 525–534 (2012).

11. Hamm, S. et al. On the mechanism of demethylation of 5-methylcytosine in DNA. *Bioorg. Med. Chem. Lett.* **18**, 1046–1049 (2008).
12. Mosammaparast, N. & Shi, Y. Reversal of histone methylation: biochemical and molecular mechanisms of histone demethylases. *Annu. Rev. Biochem.* **79**, 155–179 (2010).
13. Burgos-Barragan, G. et al. Mammals divert endogenous genotoxic formaldehyde into one-carbon metabolism. *Nature* **548**, 549–554 (2017).
14. Lyapina, M. Allergic contact dermatitis from formaldehyde exposure. *Contact Dermat.* **27**, 27–36 (1992).
15. Rea, W. J. *Chemical Sensitivity Clinical Manifestations of Pollutant Overload*. (CRC Lewis Publishers, 1996).
16. de Oliveira Ramos, C et al. The exposure to formaldehyde causes renal dysfunction, inflammation and redox imbalance in rats. *Exp. Toxicol. Pathol.* **69**, 367–372 (2017).
17. Pontel, L. B. et al. Endogenous formaldehyde is a hematopoietic stem cell genotoxin and metabolic carcinogen. *Mol. Cell* **60**, 177–188 (2015).
18. Hopkinson, R. J. & Schofield, C. J. Deciphering functions of intracellular formaldehyde: linking cancer and aldehyde metabolism. *Biochemistry* **57**, 904–906 (2018).
19. McGhee, J. D. & von Hippel, P. H. Formaldehyde as a probe of DNA structure. 3. Equilibrium denaturation of DNA and synthetic polynucleotides. *Biochemistry* **16**, 3267–3276 (1977).
20. Lu, K. et al. Structural characterization of formaldehyde-induced cross-links between amino acids and deoxynucleosides and their oligomers. *J. Am. Chem. Soc.* **132**, 3388–3399 (2010).
21. López-mirabal, H. R. & Winther, J. R. Redox characteristics of the eukaryotic cytosol. *Biochim. Biophys. Acta* **1783**, 629–640 (2008).
22. Shishodia, S. et al. NMR analyses on N-hydroxymethylated nucleobases—implications for formaldehyde toxicity and nucleic acid demethylases. *Org. Biomol. Chem.* **16**, 4021–4032 (2018).
23. McGhee, J. D. & Hippel, P. H. Von Formaldehyde as a probe of DNA structure 4 mechanism of the initial reaction of formaldehyde with DNA. *Biochemistry* **16**, 3276–3293 (1977).
24. Brown, J. H. The formal titration of bacteriological media. *J. Bacteriol.* **8**, 245–267 (1922).
25. Wadsworth, A. & Pangborn, M. C. Formaldehyde reaction with amino acids. *J. Biol. Chem.* **116**, 423–436 (1936).
26. Smith, A. K., Handler, P. H. & Mrgudich, J. N. The reaction of formaldehyde with amino acids: X-ray diffraction patterns. *J. Phys. Chem.* **44**, 874–880 (1940).
27. Middlebrook, W. R. & Phillips, H. The action of formaldehyde on the cystine disulphide linkages in wool. *Biochem. J.* **36**, 294–302 (1942).
28. French, D. & Edsall, J. T. *Advances in Protein Chemistry: The Reactions of Formaldehyde with Amino Acids and Proteins*. (Academic Press Inc., 1945).
29. Fraenkel-Conrat, H. & Olcott, H. S. Reaction of formaldehyde with proteins. II. Participation of the guanidyl groups and evidence of crosslinking. *J. Am. Chem. Soc.* **68**, 34–37 (1946).
30. Fraenkel-Conrat, H. & Olcott, H. S. The reaction of formaldehyde with proteins. V. Cross-linking between amino and primary amide or guanidyl groups. *J. Am. Chem. Soc.* **70**, 2673–2684 (1948).
31. Middlebrook, W. R. The irreversible combination of formaldehyde with proteins. *Biochem. J.* **44**, 17–23 (1949).
32. Guthe, K. F. & Arbor, A. The formaldehyde-hemoglobin reaction. *J. Biol. Chem.* **234**, 3169–3174 (1969).
33. Myers, J. S. & Hardman, J. K. Formaldehyde-induced in the a subunit *Escherichia coli* tryptophan synthetase. *J. Biol. Chem.* **246**, 3863–3869 (1971).
34. Takahashi, K. The reactions of phenylglyoxal and related reagents with amino acids. *J. Biochem.* **81**, 395–402 (1977).
35. Kitamoto, Y. & Maeda, H. Reevaluation of the reaction of formaldehyde at low concentration with amino acids. *J. Biochem.* **87**, 1519–1530 (1980).
36. Tome, D., Naulet, N. & Kozłowski, A. 13C-N.M.R. characterization of formaldehyde bonds in model mixtures and proteins containing lysine. *Int. J. Pept. Protein Res.* **25**, 258–266 (1985).
37. Joshi, N. S., Whitaker, L. R. & Francis, M. B. A three-component Mannich-type reaction for selective tyrosine bioconjugation. *J. Am. Chem. Soc.* **126**, 15942–15943 (2004).
38. Clarke, H. T., Gillespie, B. & Weisshaus, S. Z. The action of formaldehyde on amines and amino acids. *J. Am. Chem. Soc.* **55**, 4571–4587 (1933).
39. Fülöp, F., Mattinen, J. & Piñlaja, R. Ring-chain tautomerism in 1,3-thiazolidines. *Tetrahedron* **46**, 6545–6552 (1990).
40. Trézil, L. et al. Antagonistic reactions of arginine and lysine against formaldehyde and their relation to cell proliferation, apoptosis, folate cycle and photosynthesis. *Mol. Cell. Biochem.* **244**, 167–176 (2003).
41. Strack, E., Kunz, W. & Noack, R. Zur bestimmung von 4-carboxy-1,3-thiazan und von methionin. *Biol. Chem.* **313**, 201 (1958).
42. Conti, S., Cossu, S., Giacomelli, G. & Falorni, M. Chiral ligands containing heteroatoms: 13. 1 optically active 4-(2'-pyridyl)-1,3-oxazolines: an improved synthesis of 2-(2'-pyridyl)-2-aminoalcohols. *Tetrahedron* **50**, 13493–13500 (1990).
43. Mason, R. P., Sanders, J. K. M. & Gidley, M. J. NMR visualization of free asparagine in potato tissue using adduct formation with [13C]Formaldehyde. *Phytochemistry* **25**, 1567–1571 (1986).
44. Boreisha, I. K. et al. Spinacine—an inhibitor of γ -aminobutyric acid uptake. *Pharm. Chem. J.* **22**, 15–17 (1988).
45. Remelli, M., Pulidori, F., Guerrini, R. & Bertolasi, V. Synthesis of spinacine and spinacine derivatives: crystal and molecular structures of N π -hydroxymethyl spinacine and Na-methyl spinaceamine. *J. Chem. Crystallogr.* **27**, 507–513 (1991).
46. Navakauskiene, R. et al. Histone demethylating agents as potential S-adenosyl-l-methionine-competitors. *MedChemComm* **7**, 1245–1255 (2016).
47. Dale, J. & Sigvartsen, T. Condensation of alkanediamines with formaldehyde; intra molecular disproportionation of N-hydroxymethyl groups into N-methyl and N-formal groups. *Acta Chem. Scand.* **45**, 1064–1070 (1991).
48. Markolovic, S., Wilkins, S. E. & Schofield, C. J. Protein hydroxylation catalyzed by 2-oxoglutarate-dependent oxygenases. *J. Biol. Chem.* **290**, 20712–20722 (2015).
49. Metz, B. et al. Identification of formaldehyde-induced modifications in proteins. *J. Biol. Chem.* **279**, 6235–6243 (2004).
50. McDonough, M. A. et al. Structures of two kinetic intermediates reveal species specificity of penicillin-binding proteins. *J. Mol. Biol.* **322**, 111–122 (2002).
51. Luo, W., Li, H., Zhang, Y. & Ang, C. Y. Determination of formaldehyde in blood plasma by high-performance liquid chromatography with fluorescence detection. *J. Chromatogr. B Biomed. Sci. Appl.* **753**, 253–257 (2001).
52. Hopkinson, R. J., Barlow, P. S., Schofield, C. J. & Claridge, T. D. W. Studies on the reaction of glutathione and formaldehyde using NMR. *Org. Biomol. Chem.* **8**, 4915–4920 (2010).
53. Sanders, J. K. & Gidley, M. J. Reductive methylation of proteins with sodium cyanoborohydride. *Biochem. J.* **203**, 331–334 (1982).
54. Henry, L., Leung, I. K. H., Claridge, T. D. W. & Schofield, C. J. γ -Butyrobetaine hydroxylase catalyses a Stevens type rearrangement. *Bioorg. Med. Chem. Lett.* **22**, 4975–4978 (2012).
55. Bateman, R. & Shokat, K. M. Glutathione traps formaldehyde by formation of a bicyclo [4,4,1] undecane adduct. *Org. Biomol. Chem.* **5**, 3363–3367 (2007).
56. Chakraborty, A. A. et al. Histone demethylase KDM6A directly senses oxygen to control chromatin and cell fate. *Science* **363**, 1217–1222 (2019).
57. Batie, M. et al. Hypoxia induces rapid changes to histone methylation and reprograms chromatin. *Science* **363**, 1222–1226 (2019).
58. McGhee, J. D. & Von Hippel, P. H. Formaldehyde as a Probe of DNA Structure. I. Reaction with exocyclic amino groups of DNA bases. *Biochemistry* **14**, 1281–1296 (1975).
59. Mannich, C. & Krösche, W. Ueber ein kondensationsprodukt aus formaldehyd, ammoniak und antipyrin. *Arch. Pharm.* **250**, 647–667 (1912).
60. Stingle, J. et al. Mechanism and regulation of DNA-protein crosslink repair by the DNA-dependent metalloprotease SPRTN. *Mol. Cell* **64**, 688–703 (2016).
61. Friede, H. A., Campoli-Richards, D. M. & Goa, K. L. Sultamicillin. *Drugs* **37**, 491–522 (1989).
62. Swenberg, J. A. et al. Formaldehyde carcinogenicity research: 30 years and counting for mode of action, epidemiology, and cancer risk assessment. *Toxicol. Pathol.* **41**, 181–189 (2013).
63. McNamara, C. R. et al. TRPA1 mediates formalin-induced pain. *Proc. Natl Acad. Sci.* **104**, 13525–13530 (2007).
64. Wang, L., Cvetkov, T. L., Chance, M. R. & Moiseenkova-bell, V. Y. Identification of in vivo disulfide conformation of TRPA1 ion channel. *J. Biol. Chem.* **287**, 6169–6176 (2012).
65. Camalier, R. F., Kaur, G., Hollingshead, M., Holbeck, S. & Sinovi, V. S. Biological evaluation of tubulysin A: a potential anticancer and antiangiogenic natural product. *Biochem. J.* **242**, 235–242 (2006).
66. Abdulla, S. A., Zailina, H. & Abu Bakar, F. Health risk assessment of adults consuming commercial fish contaminated with formaldehyde. *Food Public Health* **3**, 52–58 (2013).
67. Jagtap, R. M., Thorat, S. H., Gonnade, R. G., Khan, A. A. & Pardeshi, S. K. X-ray crystal structures and anti-breast cancer property of 3-tert-butoxycarbonyl-2-arylthiazolidine-4-carboxylic acids. *New J. Chem.* **42**, 1078–1086 (2018).
68. Barve, A. et al. Differences in heterocycle basicity distinguish homocysteine from cysteine using aldehyde-bearing fluorophores. *Chem. Commun.* **50**, 8219–8222 (2014).
69. Ikonaka, M., Matsumoto, J. & Nishimoto, Y. A concise synthesis of (2S, 3S)-BocAHPBA and (R)-BocDMTA, chiral building blocks for peptide-mimetic HIV protease inhibitors. *Tetrahedron* **13**, 1201–1208 (2002).
70. Vishnumaya, M. R. & Singh, V. K. Highly efficient small organic molecules for enantioselective direct aldol reaction in organic and aqueous media. *J. Org. Chem.* **74**, 4289–4297 (2009).

Acknowledgements

We thank the Wellcome Trust and Cancer Research UK for funding. We thank the Engineering and Physical Sciences Research Council for a studentship to JJAGK via the Centre for Doctoral Training in Synthesis for Biology and Medicine (EP/L015838/1), and a Clarendon Scholarship.

Author contributions

J.J.A.G.K. performed the experimental work, developed the project, and undertook the data interpretation. R.J.H. and T.D.W.C. contributed to the project design and data interpretation. C.J.S. conceived and directed the project. All authors contributed to preparing the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s42004-019-0224-2>.

Correspondence and requests for materials should be addressed to R.J.H., C.J.S. or T.D.W.C.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019