

Diselenide Crosslinks for Enhanced and Simplified Oxidative Protein Folding

Reem Mousa

The Hebrew University

Taghreed Hidmi

The Hebrew University

Sergei Pomyalov

The Hebrew University

Shifra Lansky

The Hebrew University

Lareen Khouri

The Hebrew University

Debbie Shalev

The Hebrew University

Gil Shoham

The Hebrew University

Norman Metanis (✉ metanis@mail.huji.ac.il)

The Hebrew University

Article

Keywords:

Posted Date: September 16th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-73016/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Communications Chemistry on March 5th, 2021. See the published version at <https://doi.org/10.1038/s42004-021-00463-9>.

Diselenide Crosslinks for Enhanced and Simplified Oxidative Protein Folding

Reem Mousa,^[a] Taghreed Hidmi,^[a] Sergei Pomyalov,^[a] Shifra Lansky,^[a] Lareen Khouri,^[a]
Deborah E. Shalev,^[b,c] Gil Shoham^{[a]*}, and Norman Metanis^{[a]*}

[a] Institute of Chemistry, The Hebrew University of Jerusalem, Edmond J. Safra, Givat Ram,
Jerusalem 9190401, Israel

[b, c] Department of Pharmaceutical Engineering, Azrieli College of Engineering Jerusalem and
Wolfson Centre for Applied Structural Biology, The Hebrew University of Jerusalem,
Edmond J. Safra, Givat Ram, Jerusalem 9190401, Israel

Corresponding authors E-mail: Gil2@mail.huji.ac.il and Metanis@mail.huji.ac.il

Abstract. The oxidative folding of proteins has been studied for over sixty years, providing critical insight into protein folding mechanisms. Hirudin, the most potent natural inhibitor of thrombin, is a 65-residue protein with three disulfide bonds, and is viewed as a folding model for a wide range of disulfide-rich proteins. Hirudin's folding pathway is notorious for its highly heterogeneous intermediates and scrambled isomers, which plague its folding rate and yield *in vitro*. Aiming to overcome these limitations, we undertook a systematic investigation of diselenide bridges at native and non-native positions and investigated their effect on hirudin's folding, structure and activity. Our studies demonstrated that, regardless of the specific positions of these substitutions, the diselenide crosslinks enhanced the folding rate and yield of the corresponding hirudin analogs, while reducing the complexity and heterogeneity of the process. Moreover, crystal structure analysis confirmed that the diselenide substitutions maintained the overall structure of the protein and left the function virtually unchanged. The choice of hirudin as a study model has implications beyond its specific folding mechanism, demonstrating the high potential of diselenide substitutions in the design, preparation and characterization of disulfide-rich proteins.

Introduction

The *in vitro* oxidative folding of many proteins is rather sluggish process, often making it a major bottleneck in the preparation of therapeutic, disulfide-rich proteins.^{1,2} During this process, a reduced/unfolded polypeptide undergoes both conformational folding and intermediate disulfide bond formation, eventually reaching the native state (N) of the protein with its final disulfide crosslinks.³⁻⁵ On the other hand, the slow kinetics of disulfide bond formation and shuffling enable chemical trapping, isolation and further characterization of the folding intermediates,^{3,6} allowing an important in-depth elucidation of the folding process. Extensive studies of such processes have shown divergent folding pathways that differ mainly in the heterogeneity of the intermediates.³⁻⁵ At the ends of this wide spectrum of processes, there are two extreme mechanism models, represented by bovine pancreatic trypsin inhibitor (BPTI) and hirudin.³⁻⁵ BPTI folding, which represents one extreme of *in vitro* folding, is characterized by the predominance of a finite number of distinct intermediates. These intermediates have native-like structures, which assume one, and then two, native disulfide bonds, before reaching the final conformational state of the protein with its three native disulfide bonds.⁷⁻⁹ The other extreme is hirudin, a small protein isolated from the leech *Hirudo medicinalis*¹⁰, which is known to be the most potent natural inhibitor of thrombin.^{11,12} Hirudin contains two major parts, a compact N-terminal domain, held by three disulfide crosslinks, and an extended, practically unstructured, C-terminal peptide (Fig. 1a).¹³ Unlike BPTI, the folding pathway of hirudin involves numerous, highly heterogeneous disulfide-bonded intermediates, generally known as a “trial and error” mechanism.^{4,14,15} Reduced hirudin folds into its final native state in two major stages. First, the protein undergoes a non-specific packing stage, in which the protein pairs random disulfides to form heterogeneous intermediate ensembles, involving a single disulfide (1-SS), two disulfide (2-SS) and three disulfide (3-SS) scrambled species (Fig. 1b). The

second, and rate-determining stage, is the disulfide reshuffling of the heterogeneous scrambled population, leading to the formation of the final structure with its three native disulfide bonds, 6–14, 16–28 and 22–39 (Fig. 1a).^{4,14,15}

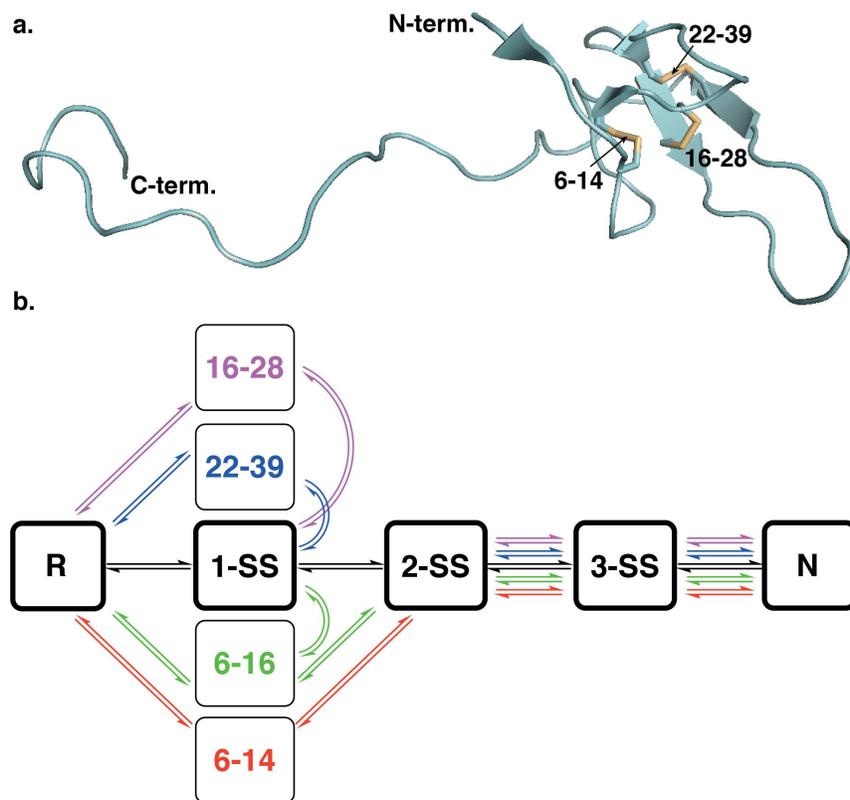


Figure 1. 3D Structure and oxidative folding mechanism of hirudin. **(a)** 3D structure of hirudin (PDB entry: 1HRT) based on crystal structure of its complex with bovine thrombin, which is not shown here. The three native disulfide bonds 6–14, 16–28, and 22–39 are shown as sticks (*yellow*); **(b)** The proposed oxidative folding mechanism of WT-Hir (in *black*) and its seleno-analogs (*coloured*).^[12,13] R and N are reduced/unfolded and native hirudin, respectively; 1-SS, 2-SS and 3-SS are ensembles of molecules with the corresponding number of disulfide bonds. The four synthetic Se-Hir analogs (*coloured*) contain a diselenide bond at the specified positions. All Se-Hir analogs folded faster than WT-Hir, and with fewer intermediates. Although all analogs formed the 2-SS and 3-SS ensembles, the 1-SS populations were mostly absent in Hir(C6U/C14U) and Hir(C6U/C16U) analogs.

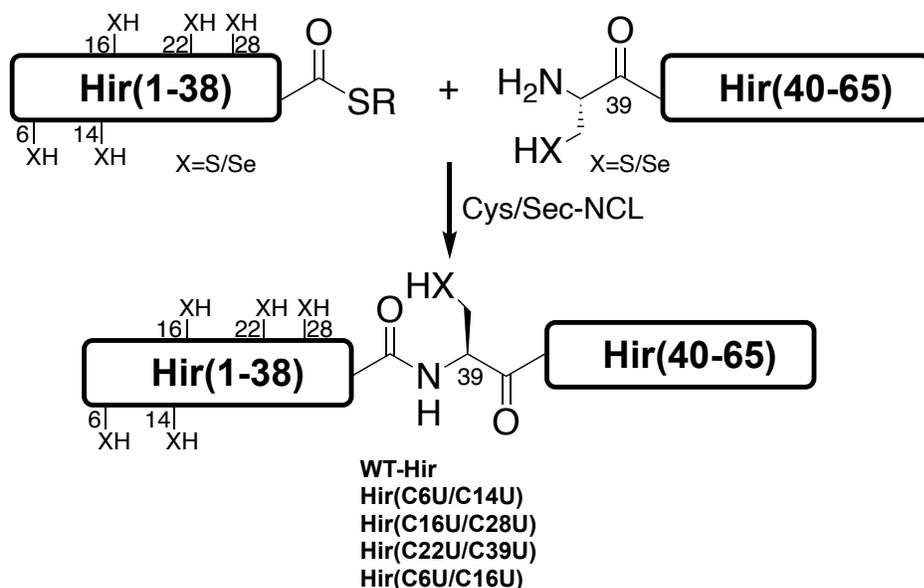
Scrambled isomer formation and the accumulation of trapped intermediates are known to decrease the efficiency of *in vitro* oxidative folding of many proteins, thereby limiting their folding rate and yield.¹⁶ One of the potential approaches to overcome such limitations involves the use of selenium-containing molecules, which can act as inter/intramolecular catalysts in disulfide exchange reactions.¹⁷⁻²⁴ A similar approach involves cysteine to selenocysteine (Sec, U) substitutions, which has been used to study the oxidative folding of several proteins.^{21,25-31} In fact, this isosteric replacement strategy was applied to native and non-native disulfides in hirudin's opposite model, BPTI.³¹⁻³⁴ In that study, the seleno-BPTI (Se-BPTI) analogs were found to fold correctly into the native state, some of which were demonstrated to bypass the common trapped intermediates of the folding pathway.³³ These improved folding results of the Se-BPTI case study motivated us to investigate the effect of diselenide substitution on the folding pathway of hirudin, which has a drastically different folding pattern. Given the more chaotic folding pathway of hirudin, we speculated that the introduction of diselenide crosslinks within the hirudin protein would decrease the distribution of folding intermediates, and thereby improve the overall folding rate. That is, by minimizing the random disulfide pairings inherent to the native folding pathway, it was expected that the productive folding routes would be enhanced over the unproductive ones, leading to a more efficient folding process.

In the present work, following the above rationale, we prepared and characterized four seleno-hirudin (Se-Hir) analogs in order to study the effect of Sec substitutions on the folding, structure and biological activity of the protein. Three of the analogs contained diselenides at the native crosslinks, 6–14, 16–28 and 22–39, while the fourth analog introduced a diselenide bond at a non-native crosslink, 6–16. Overall, our results demonstrate that disulfide-to-diselenide substitutions of the native disulfide crosslinks improve the folding efficiency to the native state, significantly

minimizing the formation of non-productive intermediates. Remarkably, this was also the case with the non-native diselenide-containing analog 6–16. Two of the Se-Hir analogs studied here indicated a simpler folding pathway, where the previously common 1-SS intermediate ensembles were largely not observed (*vide infra*). Interestingly, the overall 3D structure of all the Se-Hir analogs studied here remained almost unchanged, as confirmed unequivocally by crystal structure studies of the corresponding Se-Hir-thrombin complexes. While two of the Se-Hir analogs showed thrombin inhibitory activity very similar to WT-Hir, two other analogs exhibited a slightly lower inhibitory activity, probably as a result of small local conformational changes around the substituted selenium atoms, mainly at positions 6, 14 and/or 16.

Results and discussion

The current study began with the full chemical synthesis of wild-type hirudin (WT-Hir) and the four Se-Hir analogs described above. All five proteins were chemically prepared from two segments and a single native chemical ligation (NCL)^{35,36} reaction at the Gln38-Cys39 junction (Scheme 1).³⁷ The corresponding C-terminal peptides Hir(39-65), containing Cys/Sec39, were synthesized by Fmoc-SPPS,^{38,39} purified and characterized by HPLC and ESI-MS (Fig. S1 and S2). The Hir(1-38)-COSR N-terminal peptides were similarly prepared by Fmoc-SPPS, bearing a C-terminal thioester surrogate^{40,41} and Sec at different specific positions (6, 14, 16, 22 and/or 28) as needed (for details see SI, and Fig. S3-S7). Cys/Sec-NCL⁴²⁻⁴⁵ was then performed for the preparation of all the analogs studied (Fig. S8-S12), providing milligram quantities of the homogenous proteins. For the Se-Hir analogs this resulted in proteins with a single diselenide crosslink and four reduced thiols, while the WT-Hir protein was isolated in the fully reduced form, as confirmed by MS analysis.



Scheme 1. Chemical synthesis of WT-Hir and its seleno-analogs; Hir(C6U/C14U), Hir(C16U/C28U), Hir(C22U/C39U) and Hir(C6U/C16U) by native chemical ligation (NCL) reactions.

Once the proteins were synthesized and purified, they were subjected to oxidative folding under anaerobic conditions at pH 8.7, using oxidized glutathione (GSSG) as the oxidant (see details in the SI).¹⁵ For WT-Hir, under these conditions, the common three ensemble populations, namely 1-SS, 2-SS, and 3-SS, could be clearly observed during the folding process. These species overlapped extensively and hence appeared in the HPLC output as broad peaks (Fig. 2a), in good agreement with previous reports.^{15,46} Within 0.5 h, the 3-SS species predominated and remained stable even after 1.5 h.^{16,47} The final stage of reshuffling ended only after 7 h, yielding about 88% of the native form of WT-Hir (Table 1, Fig. 2a, Fig. S20).

For the Se-Hir analog Hir(C16U/C28U), a significantly faster folding process was observed under the same conditions, as compared to WT-Hir. More than half of the protein reached the native state within only 0.5 h, and the entire folding process was completed within 1.5 h, yielding about 90% of the native form of the protein (Table 1, Fig. 2b, Fig. S21c). The intermediates of this process

appeared as fewer and more distinct peaks, containing mainly 2- and 3-crosslinks. This pattern indicates a significant decrease in the heterogeneous populations (1-SS, 2-SS and 3-SS), where scrambled intermediates predominate and hence rearrange more rapidly to the final native state. Thus, Sec substitution at positions 16 and 28 lead to a fivefold increase in the folding rate relative to WT-Hir, with comparable overall yield, and is in good correlation with the initial predictions made above. The advantage of using Sec in oxidative folding is even more significant under aerobic conditions and in the absence of additional oxidants.^{30,33} Air oxidation of Hir(C16U/C28U) resulted in 80% overall yield of folded protein after 5 h, while under the same conditions, the folding of WT-Hir resulted in only 20% overall yield of the folded protein after 22 h (Fig. 3a and S13). These results also emphasize the importance of GSSG as a component in oxidative protein folding, as it appears to be critical not only as an oxidant but also as a catalyst for the disulfide (and diselenide) reshuffling.³⁻⁵ Additionally, it is worth noting that although Sec substitution steers the folding process to form higher ratios of productive intermediates, a small number of trapped unproductive intermediates could be observed alongside the main process, which appear to be stable even after 7 h. This indicates that scrambled isomers of this Se-Hir analog react differently towards reshuffling,¹⁶ and a small fraction becomes even more stable as a result of the Cys to Sec substitutions.

The Se-Hir analog Hir(C22U/C39U) also showed a significant improvement in folding efficiency under anaerobic conditions. About 66% of the protein reached the native state within 5 min, and the entire folding process was completed within 1.5 h, yielding about 94% of the native form of the protein (Table 1). Interestingly, a major HPLC peak was detected at 23 min, which contained mainly the 3-SS intermediate species (Fig. 2c, Fig. S21d). Although the folding yield of the Hir(C6U/C14U) proved to be similar to that of other proteins (Table 1, Fig. 2d, and Fig. S21b),

yet, this analog formed significantly fewer intermediate peaks during the folding process (Fig. 2d). These results are in good correlation with previous kinetic and computational studies, which indicated that not all the disulfide crosslinks have a similar role during the folding process of hirudin. The 6–14 crosslink has been suggested to play a critical role, since it was found to be

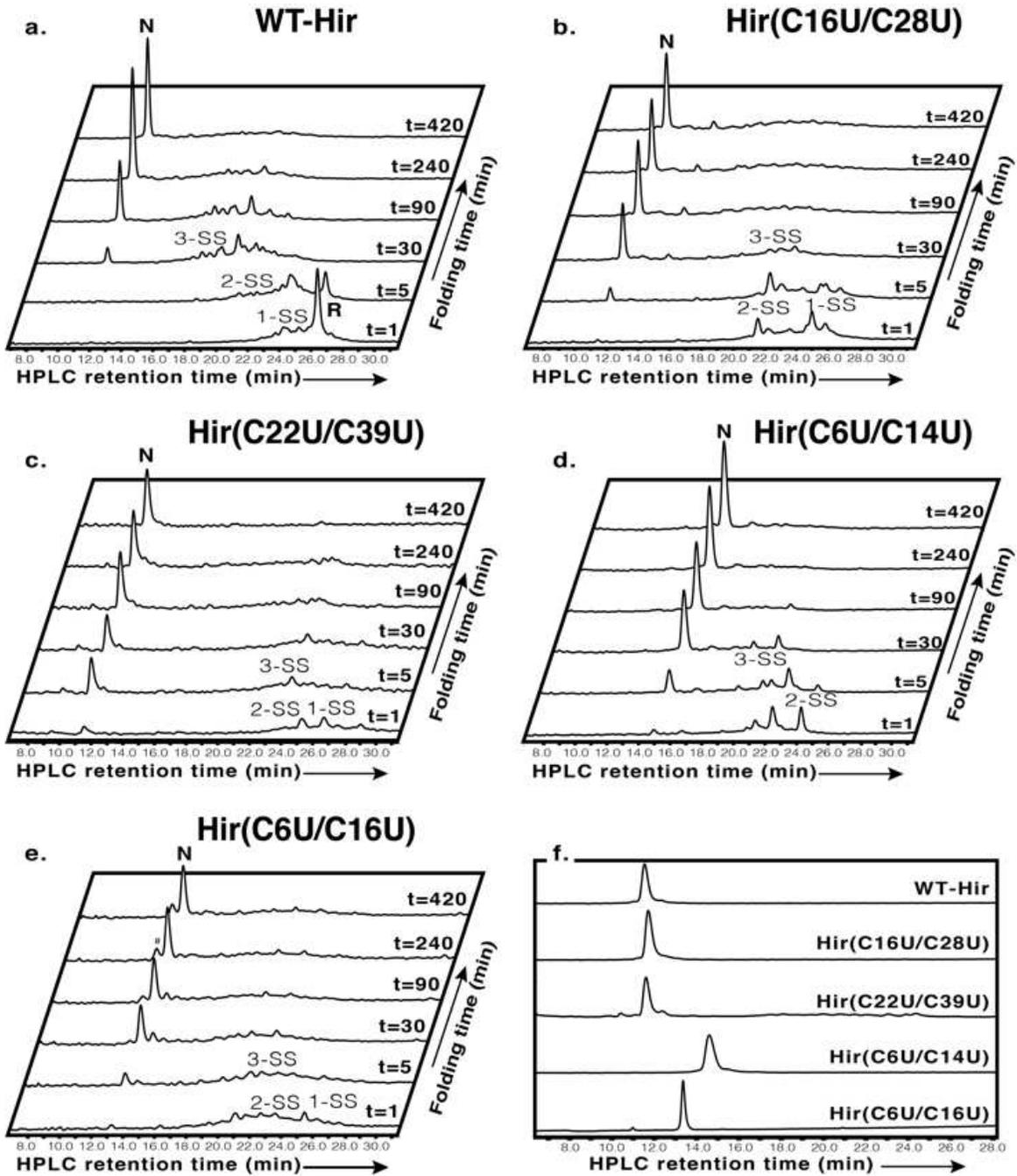


Figure 2. The anaerobic oxidative folding at pH 8.7 in the presence of 150 μ M GSSG and 30 μ M of (a) WT-Hir; (b) Hir(C16U/C28U); (c) Hir(C22U/C39U); (d) Hir(C6U/C14U); and (e) Hir(C6U/C16U). # denotes an impurity with unknown mass. 1-SS, 2-SS and 3-SS represent the number of disulfide (or diselenide) crosslinks in the intermediates. (f) Retention times of isolated native states of WT-Hir and the seleno-variants.

highly populated among the 1-SS, 2-SS, and 3-SS species in WT-Hir.^{16,48} Unlike WT-Hir, Hir(C16U/C28U) and Hir(C22U/C39U), no 1-SS species were detected during the folding process of Hir(C6U/C14U) (Fig. 2d). This can be explained by a pre-existing 6–14 diselenide crosslink, which, once present, is likely to more easily govern the formation of the remaining disulfide bonds. Thus, these current results further confirm the critical role of the 6–14 disulfide bond in the hirudin folding pathway (Fig. 1b). Despite its favorable folding rate and yield, we found that the Hir(C6U/C14U) analog displayed a lower inhibitory activity toward its natural target thrombin. The inhibitory constant K_I of this analog for thrombin was about 20-fold higher than either WT-Hir^{12,49} or the other seleno-analogs, Hir(C16U/C28U) and Hir(C22U/C39U) (Fig. 3b, Table 1). Additionally, a distinct difference was observed in the HPLC retention time of the folded form of the Hir(C6U/C14U) analog, as compared to the other seleno-variants (Fig. 2f). Although similar variations in chromatography have been reported by Alewood *et al.* of the seleno-analogs of conotoxins,^[25] the differences we observed in the retention time, combined with the lower activity against thrombin raised our concern as to whether the Hir(C6U/C14U) analog had correctly folded into the native state.

Table 1. Characterization of WT-Hir and the Se-Hir analogs.

Protein	Yield% by HPLC integration (isolated)	HPLC r.t. of N ^[a] (min)	K_I (pM) ^[b]
WT-Hir	88% (61±1%)	11.2	10.9±4.9
Hir(C16U/C28U)	90% (56±1%)	11.3	10.0±3.7
Hir(C22U/C39U)	94% (61±1%)	11.3	12.5±2.9
Hir(C6U/C14U)	95% (66±1%)	14.8	192.4±21.9
Hir(C6U/C16U)	90% (46±1%)	13.2	104.9±15.0

[a] the retention time (r.t.) of the native state (N) is shown in Fig. 2f. [b] The inhibitory activity of the different hirudin analogs with thrombin was assayed in buffer at 37 °C and using N-(*p*-Tosyl)-Gly-Pro-Arg-*p*-nitroanilide as a substrate, where the initial rate of *p*-nitroaniline formation was followed at 405 nm by UV (complete details in SI).

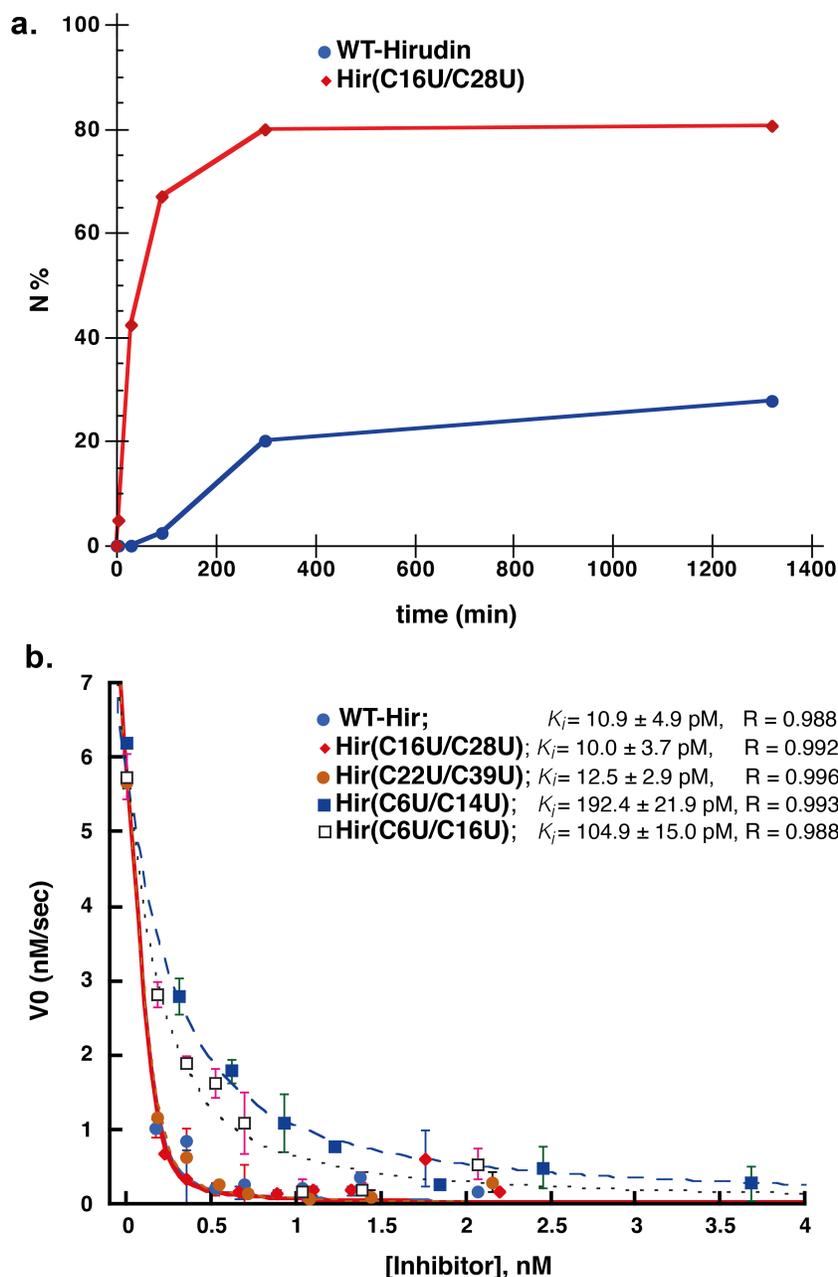


Figure 3. (a) Kinetic traces of the aerobic folding in the absence of GSSG and with 30 μ M WT-Hir or Hir(C16U/C28U) at pH 8.7. The lines connecting the data points are shown only for illustrative purposes and do not represent a data fit. (b) Inhibition activity of WT-Hir and Se-Hir analogs. The inhibition of bovine thrombin by WT-Hir (*blue circles*), Hir(C16U/C28U) (*red diamonds*), Hir(C22U/C39U) (*brown circles*), Hir(C6U/C14U) (*blue filled squares*), and Hir(C6U/C16U) (*open squares*). The data were fitted as described in the Supporting Information to give the apparent K_I values indicated.

Thus motivated to explore the folded states, we chose to examine the detailed 3D structures of the synthetic proteins. Since free hirudin proteins proved to be quite challenging to crystallize and analyze, we undertook a comprehensive structural analysis of the relevant hirudin analogs using the crystal structures of their bound complexes with thrombin.^{13,50} The three Se-Hir analogs described above were incubated with bovine thrombin to form the corresponding complexes, namely Hir(C16U/C28U)-thrombin (Complex-1), Hir(C6U/C14U)-thrombin (Complex-2) and Hir(C22U/C39U)-thrombin (Complex-3). These complexes were crystallized and their detailed 3D structures were determined and analyzed (see details in the SI). In general, all three structures were found to be quite similar to each other, and to the reported 3D structure of the WT-Hir with bovine thrombin (PDB: 1HRT, Fig. 1a).⁵¹ A global structural comparison of the bound Se-Hir in Complexes 1-3 to the bound WT-Hir in the reference structure (1HRT), gives relatively small (overall) RMSD values of 1.620, 1.476, 1.472 Å, respectively. These values, as well as a visual superposition of each of these pairs (Fig. 4) indeed demonstrate that no major global structural differences could be observed between the four structures compared here. This confirms unequivocally that, despite the replacement of a disulfide bond with a diselenide bond at different positions, all three Se-Hir analogs achieved the native folded state, adopting practically the same 3D structure.

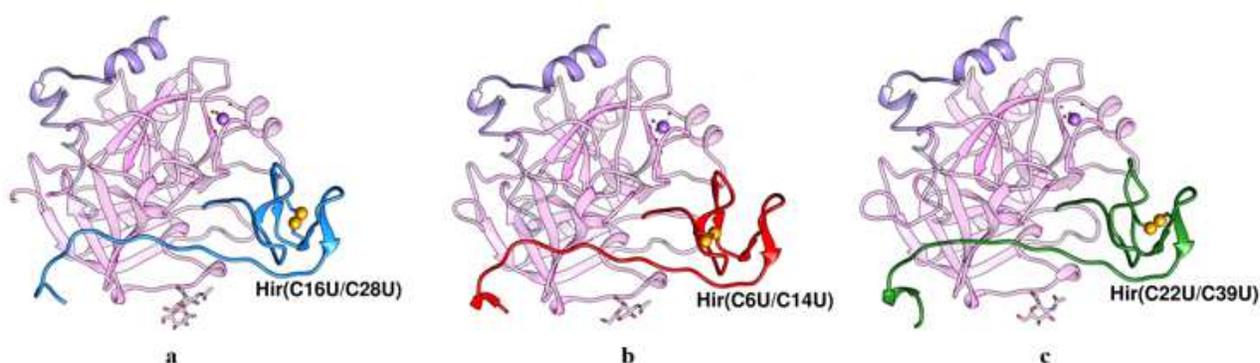


Figure 4. The overall structures of the three Se-Hir-thrombin complexes. (a) Complex-1; (b) Complex-2; (c) Complex-3. All complexes are shown in the same orientation, where the thrombin active site is at the bottom-right, opening to the right. Color codes: thrombin heavy chain (*dark purple*); thrombin light chain (*light purple*); Hir(C16U/C28U) (*blue*); Hir(C6U/C14U) (*red*); Hir(C22U/C39U) (*green*). Gold spheres represent the Se-Se bonds. PDB ID for Complex 1-3 are 7A0D, 7A0E, and 7A0F, respectively.

A closer look at the Se-Hir analogs within these complexes, and in comparison to WT-Hir, confirmed that most of the intramolecular hydrogen bonds have been closely maintained, especially the tight cluster of H-bonds located generally around the 6–14 disulfide (or diselenide) crosslink. A superposition of the Se-Hir structures in Complex-1 and Complex-2 (Table S3 in SI), demonstrates that this central H-bonding cluster remains practically unchanged, and emphasizes its central role in the stabilization of the core structure of hirudin. Especially critical within this cluster are the interactions made by Lys47 of hirudin, which plays a key role in positioning the hirudin N-terminal and C-terminal peptide segments in their specific orientations. These results reinforce previous studies, which showed that the hydrogen bonds formed by the sidechain amino group of Lys47 with the backbone carbonyl of Asp5 and the hydroxyl sidechain of Thr4 help to place the otherwise flexible N-terminal of hirudin in the active-site cleft of thrombin.⁵⁰ Interestingly, only small local conformational differences are apparent around the three original

disulfide bonds that have been substituted with diselenides, perhaps as a result of changes in the bond length, dihedral angles and atom size (Fig. S15). These local changes appear to be critical in the case of the Hir(C6U/C14U) analog, probably due to the critical location of the 6–14 crosslink within the central H-bonding cluster (Table S2). This was further supported by 2D ¹H-NMR COSY, in which some deviations in chemical shift by NMR were observed in the region around Thr4, Gln11, Leu13, Leu15 and Cys16 (Fig. S19).

Along the rationale outlined above, it was interesting to investigate how a non-native diselenide crosslink influenced the hirudin folding process. Previous studies on apamin and BPTI, where non-native diselenide substitutions were introduced, showed contradictory results. While a non-native diselenide substitution impeded the formation of the native species in the short polypeptide apamin;²⁶ a BPTI analog folded neatly into its native state with one disulfide and two selenylsulfide crosslinks.³³ In the case of hirudin, we selected the non-native crosslink 6–16 due to its predominance in 1-SS and 2-SS ensembles (*vide supra*).⁴⁸ Our experiment demonstrated that the Hir(C6U/C16U) analog folded in a similar pathway to WT-Hir, forming a heterogeneous mixture of species, mainly the 2-SS and 3-SS intermediates, but rapidly converted into the native state, with one disulfide and two selenylsulfide bonds, with a yield similar to WT-Hir (Table 1, Fig. 2e, and Fig. S21e). This analog showed a 10-fold reduction in its inhibitory activity toward thrombin (Table 1, Fig. 3b), similar to Hir(C6U/C14U), supporting our previous interpretation that a fine-tuned structural perturbation around Cys/Sec at 6, 14 and/or 16 are likely to be responsible for the decreased activity observed for these two analogs (*vide supra*).

Conclusions

In summary, we studied the effect of diselenide substitutions on the folding of the model protein hirudin. We show that an energetically favored diselenide crosslink formation reduces the complexity and heterogeneity of the folding pathway of hirudin and thereby enhances the folding rate and its efficiency. Interestingly, in the case of the native 6–14 and non-native 6–16 diselenide substitutions, the 1-SS populations were mostly absent, showing that the early stages of hirudin folding could be bypassed entirely^{16,48} and still giving the native state. However, these two diselenide substitutions slightly reduced the inhibitory activity of the corresponding hirudin analogs, likely due to subtle structural perturbations surrounding the Cys/Sec6,14 and/or 16 residues. The substitution of the other two native crosslinks with diselenides had practically no effect on the 3D structure and inhibitory activity of these Se-Hir analogs, yet still demonstrated the general improvement in folding efficiency observed for all the Se-Hir analogs. These results indicate that although disulfide-to-diselenide substitution proved to be generally advantageous for the preparation of hirudin analogs, as in other proteins, the exact location of the replaced crosslink could be important to fully retain its inhibitory activity.

The choice of hirudin as model in this study seems to have implications beyond its specific folding mechanism. Unlike BPTI, the defining characteristics of hirudin lies in the complex route it follows to achieve its native fold, and this seemingly chaotic pathway is shared by many other proteins.³⁻⁵ Furthermore, many proteins adopt a folding mechanism that lies somewhere on the spectrum between these two opposite folding models.⁴ We found that the substitution of native/non-native disulfide with diselenide bonds typically enhances protein folding and yield, preserves protein structure, and often confers no significant effect on function. Based on these results, the current study of the seleno-hirudin analogs extends the previous studies on

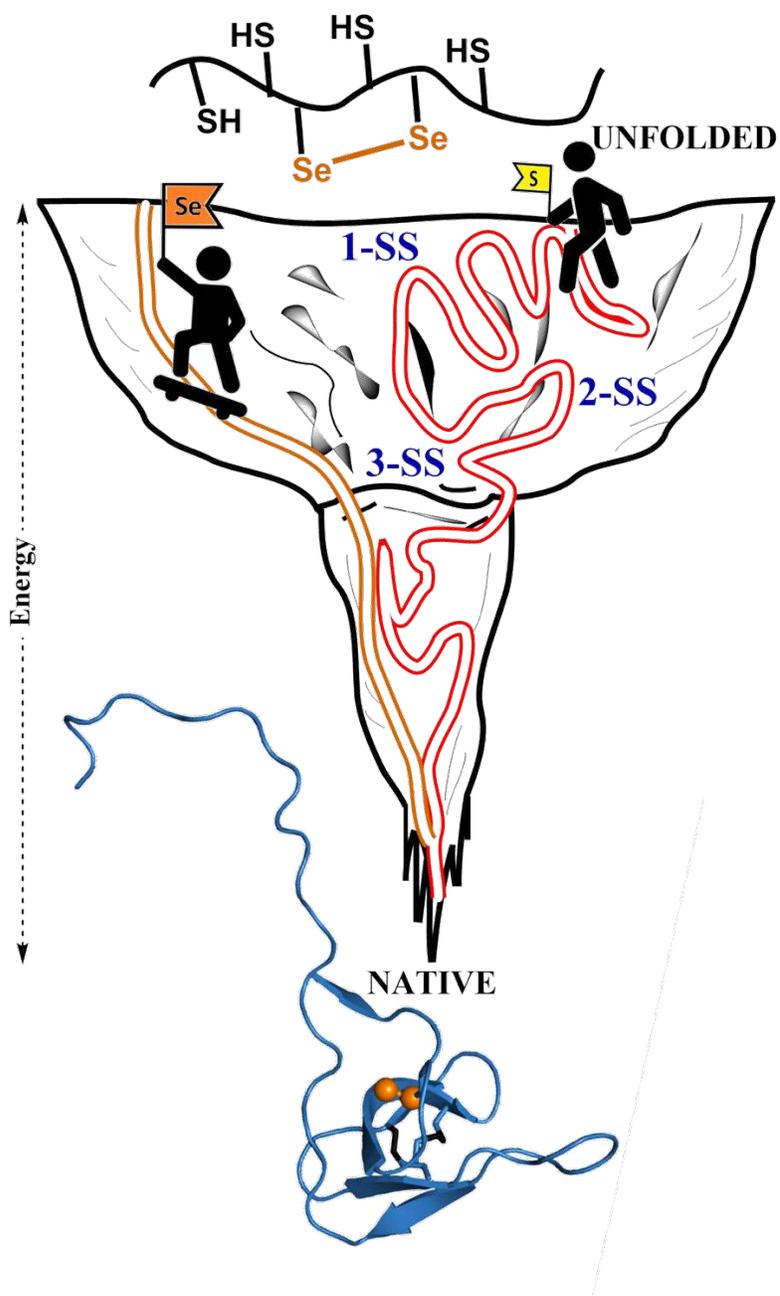
selenocysteine-containing analogs of BPTI,^{33,34} conotoxins,²⁷⁻³⁰ and insulin,⁵²⁻⁵⁴ which together demonstrate the high potential of disulfide-to-diselenide substitutions in basic and applied research. In addition to its obvious advantage in the study and preparation of disulfide-rich proteins, such an approach may also prove useful for a wide range of applications, including the rational improvement of the stability, activity and specificity of this unique family of proteins.

Acknowledgements

We thank Mrs. Ricki Notis Dardashti for input on the manuscript. N.M. acknowledges the financial support of Israel Science Foundation (1072/14 and 783/18) and ICRF Acceleration Grant. G. S. acknowledges the support of both the Israel Science Foundation (1905/15) and The Israeli Ministry of Science (3-12484/15). R. M. acknowledges the support of the VATAT scholarship for Arab students. T.H. is grateful to the Neubauer Fellowship for Arab Ph.D. students.

Keywords: Oxidative protein folding • disulphide-rich proteins • chemical protein synthesis • selenocysteine • selenoproteins

TOC Figure



Bibliography

- 1 Tombling, B. J., Wang, C. N. K. & Craik, D. J. EGF-like and Other Disulfide-rich Microdomains as Therapeutic Scaffolds. *Angew. Chem. Int. Ed.* **59**, 11218-11232 (2020).
- 2 Gongora-Benitez, M., Tulla-Puche, J. & Albericio, F. Multifaceted Roles of Disulfide Bonds. Peptides as Therapeutics. *Chem. Rev.* **114**, 901-926 (2014).
- 3 Mainathambika, B. S. & Bardwell, J. C. Disulfide-Linked Protein Folding Pathways. *Annu Rev Cell Dev Bi* **24**, 211-235 (2008).
- 4 Chang, J. Y. Diverse pathways of oxidative folding of disulfide proteins: underlying causes and folding models. *Biochemistry* **50**, 3414-3431 (2011).
- 5 Arolas, J. L., Aviles, F. X., Chang, J. Y. & Ventura, S. Folding of small disulfide-rich proteins: clarifying the puzzle. *Trends Biochem. Sci.* **31**, 292-301 (2006).
- 6 Anfinsen, C. B. Principles that govern folding of protein chains. *Science* **181**, 223-230 (1973).
- 7 Creighton, T. E. Experimental Studies of Protein Folding and Unfolding. *Prog Biophys Mol Bio* **33**, 231-297 (1978).
- 8 Weissman, J. S. & Kim, P. S. Reexamination of the folding of BPTI - predominance of native intermediates. *Science* **253**, 1386-1393 (1991).
- 9 Mousa, R., Lansky, S., Shoham, G. & Metanis, N. BPTI folding revisited: switching a disulfide into methylene thioacetal reveals a previously hidden path. *Chem. Sci.* **9**, 4814-4820 (2018).

- 10 Dodt, J., Muller, H. P., Seemuller, U. & Chang, J. Y. The Complete Amino-Acid-Sequence of Hirudin, a Thrombin Specific Inhibitor - Application of Color Carboxymethylation. *Febs Lett* **165**, 180-184 (1984).
- 11 Binnie, C. G., Erickson, B. W. & Hermans, J. Inhibition of Thrombin by Synthetic Hirudin Peptides. *Febs Lett.* **270**, 85-89 (1990).
- 12 Stone, S. R. & Hofsteenge, J. Kinetics of the inhibition of thrombin by hirudin. *Biochemistry* **25**, 4622-4628 (1986).
- 13 Rydel, T. J., Tulinsky, A., Bode, W. & Huber, R. Refined structure of the hirudin-thrombin complex. *J. Mol. Biol.* **221**, 583-601 (1991).
- 14 Chatrenet, B. & Chang, J. Y. The folding of hirudin adopts a mechanism of trial and error. *J. Biol. Chem.* **267**, 3038-3043 (1992).
- 15 Chatrenet, B. & Chang, J. Y. The disulfide folding pathway of hirudin elucidated by stop go folding experiments. *J. Biol. Chem.* **268**, 20988-20996 (1993).
- 16 Chang, J. Y. The Properties of Scrambled Hirudins. *J. Biol. Chem.* **270**, 25661-25666 (1995).
- 17 Beld, J., Woycechowsky, K. J. & Hilvert, D. Selenoglutathione: efficient oxidative protein folding by a diselenide. *Biochemistry* **46**, 5382-5390 (2007).
- 18 Beld, J., Woycechowsky, K. J. & Hilvert, D. Catalysis of oxidative protein folding by small-molecule diselenides. *Biochemistry* **47**, 6985-6987 (2008).
- 19 Beld, J., Woycechowsky, K. J. & Hilvert, D. Selenoglutathione: Efficient Oxidative Protein Folding by a Diselenide. *Biochemistry* **48**, 4662-4662 (2009).
- 20 Hondal, R. J., Marino, S. M. & Gladyshev, V. N. Selenocysteine in thiol/disulfide-like exchange reactions. *Antioxid. Redox Signal.* **18**, 1675-1689 (2013).

- 21 Mousa, R., Dardashti, R. N. & Metanis, N. Selenium and Selenocysteine in Protein Chemistry. *Angew. Chem. Int. Ed.* **56**, 15818-15827 (2017).
- 22 Mousa, R., Reddy, P. S. & Metanis, N. Chemical Protein Synthesis through Selenocysteine Chemistry. *Synlett* **28**, 1389-1393 (2017).
- 23 Reddy, P. S. & Metanis, N. Small molecule diselenide additives for in vitro oxidative protein folding. *Chem. Commun.* **52**, 3336-3339 (2016).
- 24 Metanis, N., Foletti, C., Beld, J. & Hilvert, D. Selenoglutathione-mediated rescue of kinetically trapped intermediates in oxidative protein folding. *Isr. J. Chem.* **51**, 953-959 (2011).
- 25 Pegoraro, S., Fiori, S., Rudolph-Bohner, S., Watanabe, T. X. & Moroder, L. Isomorphous replacement of cystine with selenocystine in endothelin: oxidative refolding, biological and conformational properties of [Sec(3),Sec(11),Nle(7)]-endothelin-1. *J. Mol. Biol.* **284**, 779-792 (1998).
- 26 Pegoraro, S., Fiori, S., Cramer, J., Rudolph-Bohner, S. & Moroder, L. The disulfide-coupled folding pathway of apamin as derived from diselenide-quenched analogs and intermediates. *Prot. Sci.* **8**, 1605-1613 (1999).
- 27 Muttenthaler, M. *et al.* Solving the alpha-conotoxin folding problem: efficient selenium-directed on-resin generation of more potent and stable nicotinic acetylcholine receptor antagonists. *J. Am. Chem. Soc.* **132**, 3514-3522 (2010).
- 28 de Araujo, A. D. *et al.* Total synthesis of the analgesic conotoxin MrVIB through selenocysteine-assisted folding. *Angew Chem. Int. Ed.* **50**, 6527-6529 (2011).

- 29 Gowd, K. H. *et al.* Site-specific effects of diselenide bridges on the oxidative folding of a cystine knot peptide, omega-selenoconotoxin GVIA. *Biochemistry* **49**, 2741-2752 (2010).
- 30 Steiner, A. M., Woycechowsky, K. J., Olivera, B. M. & Bulaj, G. Reagentless oxidative folding of disulfide-rich peptides catalyzed by an intramolecular diselenide. *Angew. Chem. Int. Ed.* **51**, 5580-5584 (2012).
- 31 Metanis, N. & Hilvert, D. Natural and synthetic selenoproteins. *Curr. Opin. Chem. Biol.* **22**, 27-34 (2014).
- 32 Metanis, N., Beld, J. & Hilvert, D. The chemistry of selenocysteine. *Patai's Chemistry of Functional Groups*, doi:10.1002/9780470682531.pat0582. (2011).
- 33 Metanis, N. & Hilvert, D. Strategic use of non-native diselenide bridges to steer oxidative protein folding. *Angew. Chem. Int. Ed.* **51**, 5585-5588 (2012).
- 34 Metanis, N. & Hilvert, D. Harnessing selenocysteine reactivity for oxidative protein folding. *Chem. Sci.* **6**, 322-325 (2015).
- 35 Dawson, P. E., Muir, T. W., Clark-Lewis, I. & Kent, S. B. H. Synthesis of proteins by native chemical ligation. *Science* **266**, 776-779 (1994).
- 36 Kent, S. B. H. Total chemical synthesis of proteins. *Chem Soc Rev* **38**, 338-351 (2009).
- 37 The chemical synthesis of Hirudin-P6 variant as well as the Thr43 O-glycosylated and Tyr61 sulfated analogs were previously reported by Payne and coworkers, using the Lys27-Cys28 ligation junction. Hsieh, Y. S. Y., Wijeyewickrema, L. C., Wilkinson, B. L., Pike, R. N. & Payne, R. J. Total Synthesis of Homogeneous Variants of Hirudin P6: A PostTranslationally Modified Anti- Thrombotic Leech- Derived Protein. *Angew Chem. Int. Ed.* **53**, 3947-3951 (2014).

- 38 Merrifield, R. B. Solid phase peptide synthesis .1. Synthesis of a tetrapeptide. *J. Am. Chem. Soc.* **85**, 2149-2154 (1963).
- 39 Behrendt, R., White, P. & Offer, J. Advances in Fmoc solid-phase peptide synthesis. *J. Pept. Sci.* **22**, 4-27 (2016).
- 40 Blanco-Canosa, J. B. & Dawson, P. E. An efficient Fmoc-SPPS approach for the generation of thioester peptide precursors for use in native chemical ligation. *Angew. Chem. Int. Ed.* **47**, 6851-6855 (2008).
- 41 Flood, D. T. *et al.* Leveraging the Knorr Pyrazole Synthesis for the Facile Generation of Thioester Surrogates for use in Native Chemical Ligation. *Angew Chem. Int. Ed.* **57**, 11634-11639 (2018).
- 42 Reddy, P. S., Dery, S. & Metanis, N. Chemical Synthesis of Proteins with Non-Strategically Placed Cysteines Using Selenazolidine and Selective Deselenization. *Angew. Chem. Int. Ed.* **55**, 992-995 (2016).
- 43 Hondal, R. J., Nilsson, B. L. & Raines, R. T. Selenocysteine in native chemical ligation and expressed protein ligation. *J. Am. Chem. Soc.* **123**, 5140-5141 (2001).
- 44 Gieselman, M. D., Xie, L. & van der Donk, W. A. Synthesis of a selenocysteine-containing peptide by native chemical ligation. *Org. Lett.* **3**, 1331-1334 (2001).
- 45 Quaderer, R., Sewing, A. & Hilvert, D. Selenocysteine-mediated native chemical ligation. *Helv. Chim. Acta* **84**, 1197-1206 (2001).
- 46 Chang, J. Y. Controlling the speed of hirudin folding. *Biochem. J.* **300**, 643-650 (1995).
- 47 Chang, J. Y., Schindler, P. & Chatrenet, B. The Disulfide Structures of Scrambled Hirudins. *J. Biol. Chem.* **270**, 11992-11997 (1995).

- 48 Micheletti, C., De Filippis, V., Maritan, A. & Seno, F. Elucidation of the disulfide-folding pathway of hirudin by a topology-based approach. *Proteins* **53**, 720-730 (2003).
- 49 Otto, A. & Seckler, R. Characterization, Stability and Refolding of Recombinant Hirudin. *Eur. J. Biochem.* **202**, 67-73 (1991).
- 50 Rydel, T. J. *et al.* The Structure of a Complex of Recombinant Hirudin and Human Alpha-Thrombin. *Science* **249**, 277-280 (1990).
- 51 Vitali, J. *et al.* The Structure of a Complex of Bovine Alpha-Thrombin and Recombinant Hirudin at 2.8-Å Resolution. *J. Biol. Chem.* **267**, 17670-17678 (1992).
- 52 Weil-Ktorza, O. *et al.* Substitution of an Internal Disulfide Bridge with a Diselenide Enhances both Foldability and Stability of Human Insulin. *Chem-Eur. J.* **25**, 8513-8521 (2019).
- 53 Dhayalan, B. *et al.* Reassessment of an Innovative Insulin Analogue Excludes Protracted Action yet Highlights the Distinction between External and Internal Diselenide Bridges. *Chem-Eur. J.* **26**, 4695-4700 (2020).
- 54 Arai, K. *et al.* Preparation of Selenoinsulin as a Long-Lasting Insulin Analogue. *Angew Chem. Int. Ed.* **56**, 5522-5526 (2017).

Figures

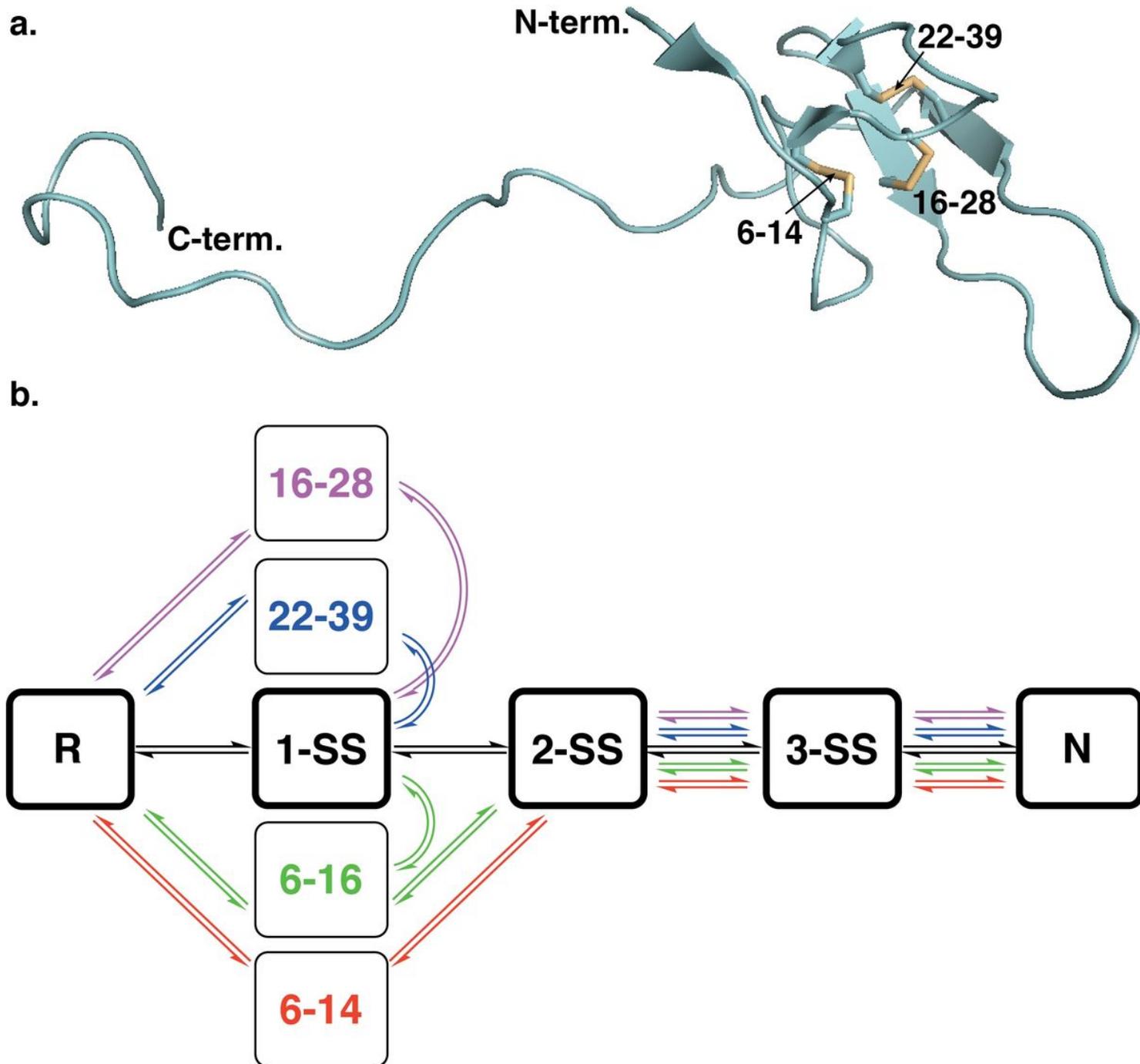


Figure 1

3D Structure and oxidative folding mechanism of hirudin. (a) 3D structure of hirudin (PDB entry: 1HRT) based on crystal structure of its complex with bovine thrombin, which is not shown here. The three native disulfide bonds 6–14, 16–28, and 22–39 are shown as sticks (yellow); (b) The proposed oxidative folding mechanism of WT-Hir (in black) and its seleno-analogs (coloured).[12,13] R and N are reduced/unfolded and native hirudin, respectively; 1-SS, 2-SS and 3-SS are ensembles of molecules with the corresponding number of disulfide bonds. The four synthetic Se-Hir analogs (coloured) contain a diselenide bond at the

specified positions. All Se-Hir analogs folded faster than WT-Hir, and with fewer intermediates. Although all analogs formed the 2-SS and 3-SS ensembles, the 1-SS populations were mostly absent in Hir(C6U/C14U) and Hir(C6U/C16U) analogs.

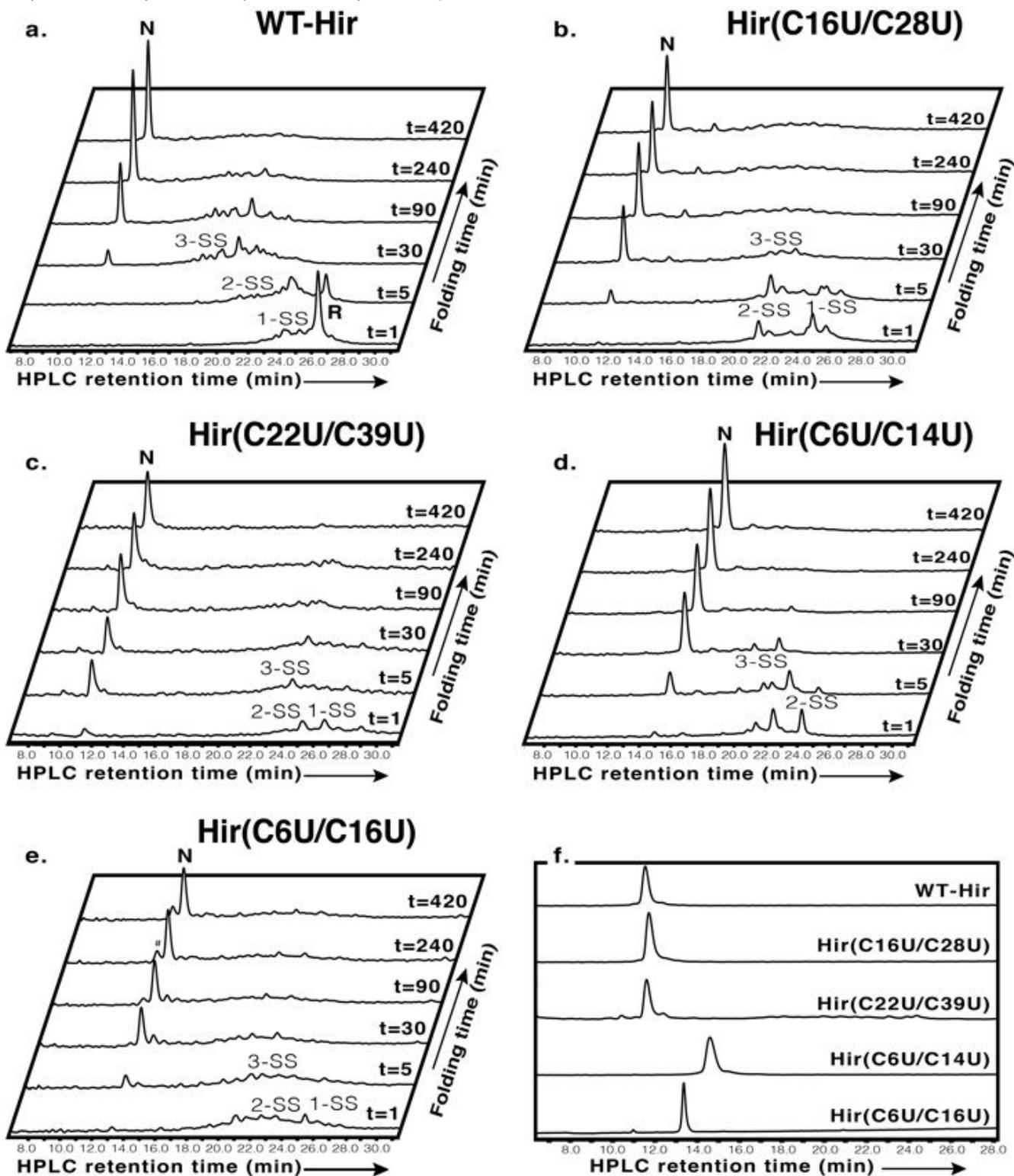


Figure 2

The anaerobic oxidative folding at pH 8.7 in the presence of 150 μ M GSSG and 30 μ M of (a) WT-Hir; (b) Hir(C16U/C28U); (c) Hir(C22U/C39U); (d) Hir(C6U/C14U); and (e) Hir(C6U/C16U). #denotes an impurity

with unknown mass. 1-SS, 2-SS and 3-SS represent the number of disulfide (or diselenide) crosslinks in the intermediates. (f) Retention times of isolated native states of WT-Hir and the seleno-variants.

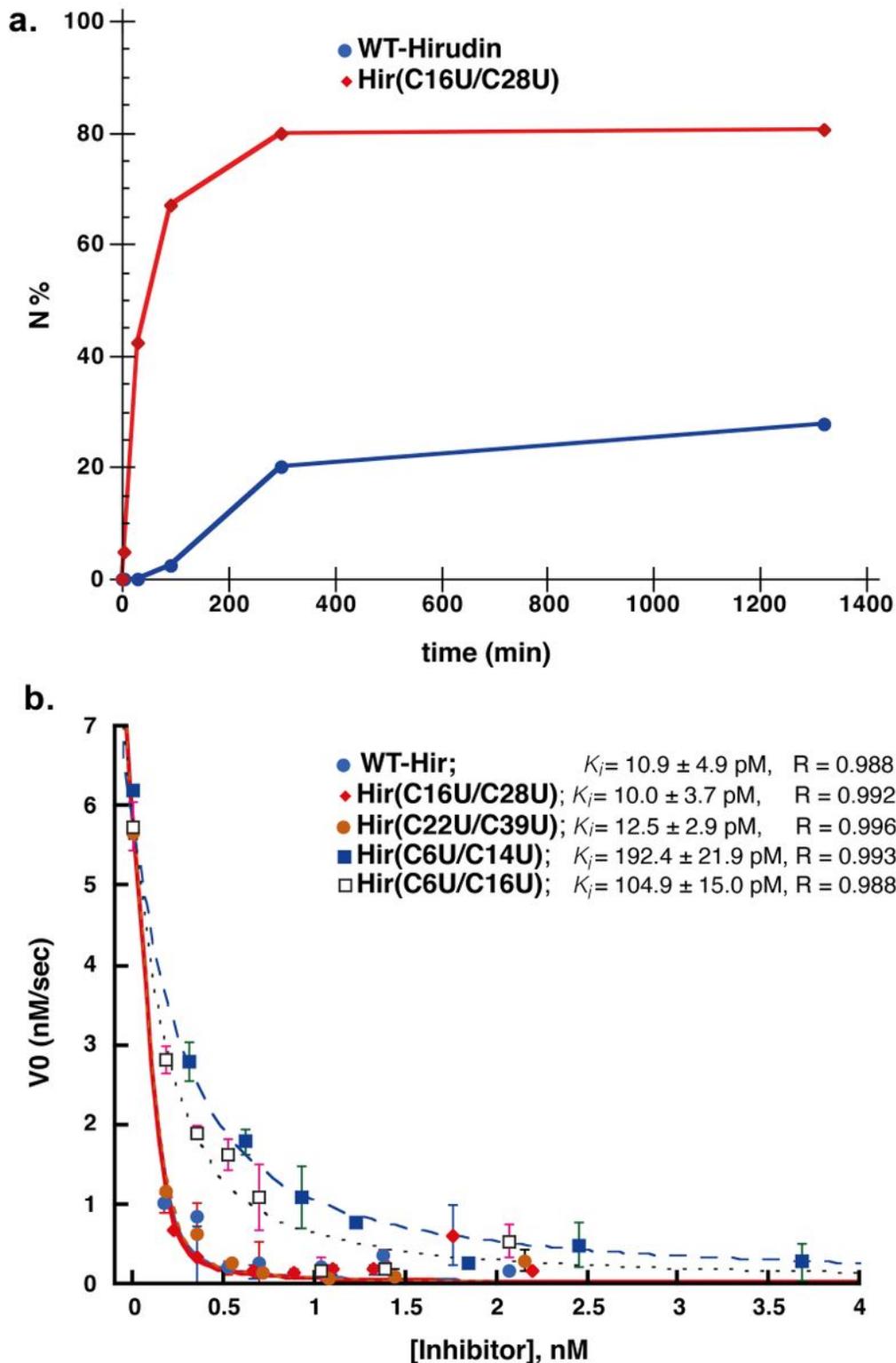


Figure 3

(a) Kinetic traces of the aerobic folding in the absence of GSSG and with 30 μ M WT-Hir or Hir(C16U/C28U) at pH 8.7. The lines connecting the data points are shown only for illustrative purposes and do not represent a data fit. (b) Inhibition activity of WT-Hir and Se-Hir analogs. The inhibition of

bovine thrombin by WT-Hir (blue circles), Hir(C16U/C28U) (red diamonds), Hir(C22U/C39U) (brown circles), Hir(C6U/C14U) (blue filled squares), and Hir(C6U/C16U) (open squares). The data were fitted as described in the Supporting Information to give the apparent KI values indicated.

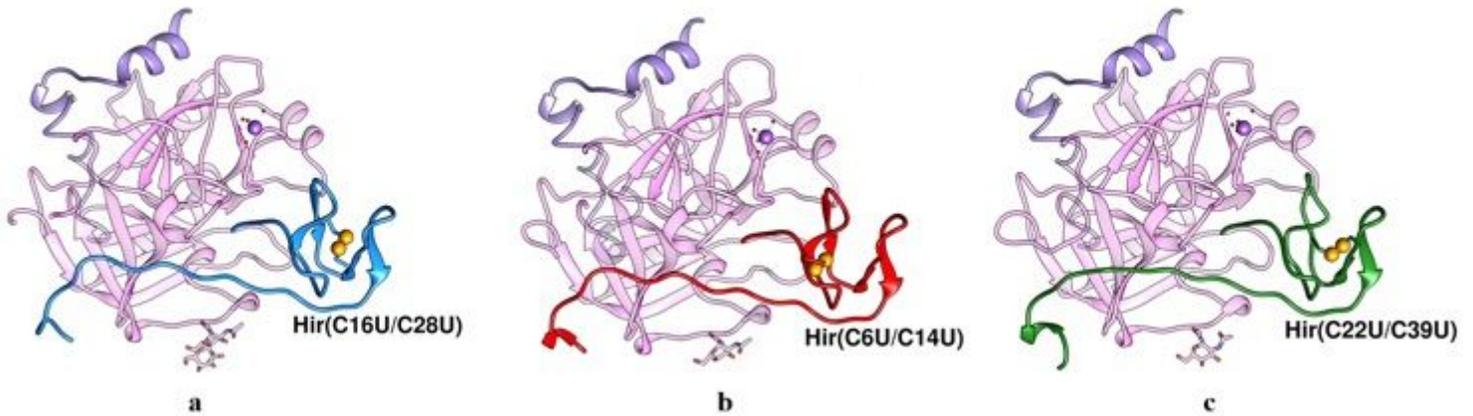


Figure 4

The overall structures of the three Se-Hir-thrombin complexes. (a) Complex-1; (b) Complex-2; (c) Complex-3. All complexes are shown in the same orientation, where the thrombin active site is at the bottom-right, opening to the right. Color codes: thrombin heavy chain (dark purple); thrombin light chain (light purple); Hir(C16U/C28U) (blue); Hir(C6U/C14U) (red); Hir(C22U/C39U) (green). Gold spheres represent the Se-Se bonds. PDB ID for Complex 1-3 are 7A0D, 7A0E, and 7A0F, respectively.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [HirudinpaperSlv21NMfinal.pdf](#)
- [7A0DComplex116U28U1.2sigma.png](#)
- [7A0FComplex322U39U1.2sigma.png](#)
- [7A0EComplex26U14U1.2sigma.png](#)
- [7A0DComplex116U28UD1292110383modelreviewP1.txt](#)
- [7A0DComplex116U28UD1292110383sfannotateP1.txt](#)
- [7A0EComplex26U14UD1292110306modelreviewP1.txt](#)
- [7A0EComplex26U14UD1292110306sfannotateP1.txt](#)
- [7A0FComplex322U39UD1292110407modelreviewP1.txt](#)
- [7A0FComplex322U39UD1292110407sfannotateP1.txt](#)
- [7A0DComplex116U28UD1292110383valreportfullannotateP1.pdf](#)
- [7A0EComplex26U14UD1292110306valreportfullannotateP1.pdf](#)
- [7A0FComplex322U39UD1292110407valreportfullannotateP1.pdf](#)