

The Role of Molecular Recognition in L-Cystine Crystal Growth Inhibition

Laura N. Poloni,[†] Zina Zhu,[†] Nelson Garcia-Vázquez,[†] Anthony Yu,[†] Misha Bhandari,[†] David Connors,[†]
Longqin Hu,[¶] Amrik Sahota,[§] Michael D. Ward,^{*,†} and Alexander G. Shtukenberg^{*,†}

[†]Department of Chemistry and the Molecular Design Institute, New York University, 100 Washington Square East, New York, NY 10003-6688, USA

[¶]Department of Medicinal Chemistry, Ernest Mario School of Pharmacy, Rutgers, The State University of New Jersey, 160 Frelinghuysen Road, Piscataway, New Jersey 08854, USA

[§]Department of Genetics, Rutgers, The State University of New Jersey, Piscataway, New Jersey 08854, USA

Table of Contents

Chart S1. L-cystine and imposters.

Chart S2. L-cysteine and imposters.

Chart S3. Therapeutic analogs of L-cystine and L-cysteine

Syntheses of Molecular Imposters

Table S1. IUPAC names of L-cystine and prospective imposters.

Table S2. Threshold additive concentration, c_{ad} , for the exclusive formation of the L-cystine tetragonal polymorph.

Figure S1. Declining concentration of L-cystine and evolution of L-cystine crystal size during bulk crystallization in the presence of L-CDME.

Figure S2. Energetically preferred orientation of adsorbed imposters on the (0001) face. (A) L-cystine, (B) L-CDME, (C) L-HCME, (D) L-CDPE, (E) L-CDMOR, (F) L-CDNMP.

Figure S3. Energetically preferred orientation of adsorbed imposters on the (10-10) face. (A) L-cystine, (B) L-CDME, (C) L-HCME, (D) L-CDPE, (E) L-CDMOR, (F) L-CDNMP.

Figure S4. The six unique “clockwise” or “+” kink sites on {10-10} steps on the {0001} L-cystine surface.

Figure S5. The six unique “clockwise” or “+” kink sites on {1-100} steps on the {10-10} L-cystine surface (left panel) and the {0001} steps on the {10-10} L-cystine surface (right panel).

Materials and Methods

Chart S1. L-cystine and imposters. Modifications to L-cystine are highlighted in red.

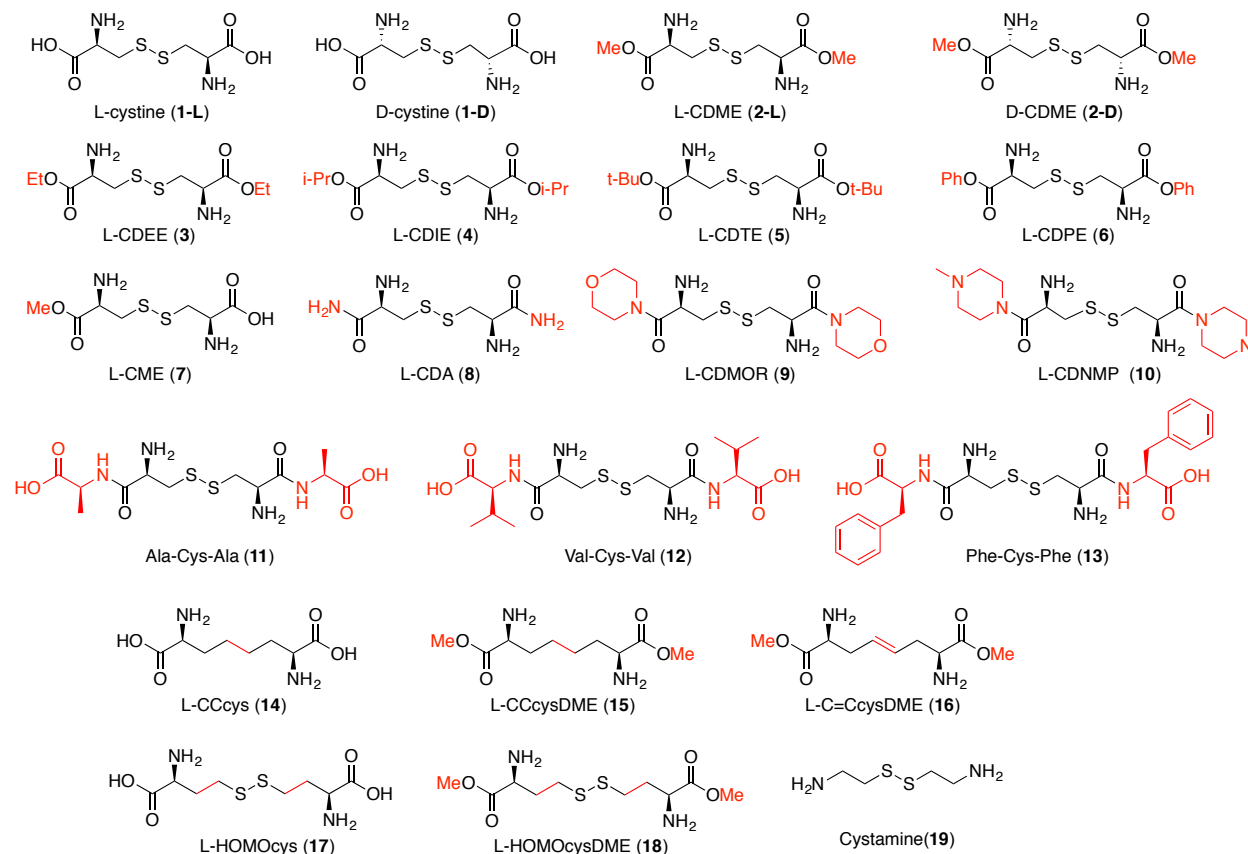


Chart S2. L-cysteine and imposters. Modifications to L-cysteine are highlighted in red.

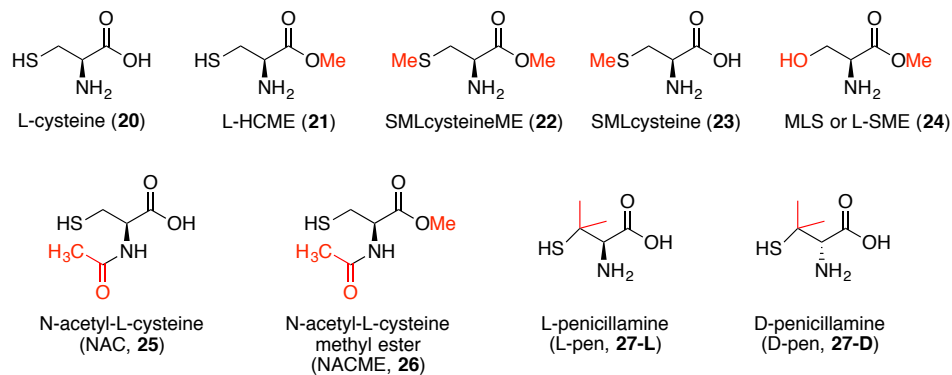


Chart S3. Therapeutic analogs of L-cystine and L-cysteine

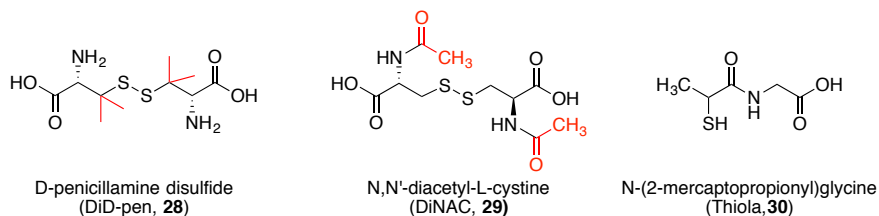


Table S1. IUPAC names of L-cystine and prospective imposters.

No.	Acronym	IUPAC name
1-L	L-cystine	(2 <i>R</i> ,2' <i>R</i>)-3,3'-disulfanediylbis(2-aminopropanoic acid)
1-D	D-cystine	(2 <i>S</i> ,2' <i>S</i>)-3,3'-disulfanediylbis(2-aminopropanoic acid)
2-L	L-CDME	dimethyl 3,3'-disulfanediyl(2 <i>R</i> ,2' <i>R</i>)-bis(2-aminopropanoate)
2-D	D-CDME	dimethyl 3,3'-disulfanediyl(2 <i>S</i> ,2' <i>S</i>)-bis(2-aminopropanoate)
3	L-CDEE	diethyl 3,3'-disulfanediyl(2 <i>R</i> ,2' <i>R</i>)-bis(2-aminopropanoate)
4	L-CDIE	diisopropyl 3,3'-disulfanediyl(2 <i>R</i> ,2' <i>R</i>)-bis(2-aminopropanoate)
5	L-CDTE	di- <i>tert</i> -butyl 3,3'-disulfanediyl(2 <i>R</i> ,2' <i>R</i>)-bis(2-aminopropanoate)
6	L-CDPE	diphenyl 3,3'-disulfanediyl(2 <i>R</i> ,2' <i>R</i>)-bis(2-aminopropanoate)
7	L-CME	<i>S</i> -(((<i>R</i>)-2-amino-3-methoxy-3-oxopropyl)thio)- <i>L</i> -cysteine
8	L-CDA	(2 <i>R</i> ,2' <i>R</i>)-3,3'-disulfanediylbis(2-aminopropanamide)
9	L-CDMOR	(2 <i>R</i> ,2' <i>R</i>)-3,3'-disulfanediylbis(2-amino-1-morpholinopropan-1-one)
10	L-CDNMP	(2 <i>R</i> ,2' <i>R</i>)-3,3'-disulfanediylbis(2-amino-1-(4-methylpiperazin-1-yl)propan-1-one)
11	Ala-Cys-Ala	(2 <i>S</i> ,2' <i>S</i>)-2,2'-(((2 <i>R</i> ,2' <i>R</i>)-3,3'-disulfanediylbis(2-aminopropanoyl))bis(azanediyl))dipropionic acid
12	Val-Cys-Val	(2 <i>S</i> ,2' <i>S</i>)-2,2'-(((2 <i>R</i> ,2' <i>R</i>)-3,3'-disulfanediylbis(2-aminopropanoyl))bis(azanediyl))bis(3-methylbutanoic acid)
13	Phe-Cys-Phe	(2 <i>S</i> ,2' <i>S</i>)-2,2'-(((2 <i>R</i> ,2' <i>R</i>)-3,3'-disulfanediylbis(2-aminopropanoyl))bis(azanediyl))bis(3-phenylpropanoic acid)
14	L-CCcys	(2 <i>S</i> ,7 <i>S</i>)-2,7-diaminooctanedioic acid
15	L-CCcysDME	dimethyl (2 <i>S</i> ,7 <i>S</i>)-2,7-diaminooctanedioate
16	L-C=CcysDME	dimethyl (2 <i>S</i> ,7 <i>S</i> , <i>E</i>)-2,7-diaminooct-4-enedioate
17	L-HOMOcys	dimethyl (2 <i>S</i> ,7 <i>S</i> , <i>E</i>)-2,7-diaminooct-4-enedioate
18	L-HOMOcysDME	dimethyl 4,4'-disulfanediyl(2 <i>S</i> ,2' <i>S</i>)-bis(2-aminobutanoate)
19	Cystamine	2,2'-disulfanediylbis(ethan-1-amine)
20	L-cysteine	<i>L</i> -cysteine
21	HCME	methyl <i>L</i> -cysteinate
22	SMLcysteineME	methyl <i>S</i> -methyl- <i>L</i> -cysteinate
23	SMLcysteine	<i>S</i> -methyl- <i>L</i> -cysteine
24	MLS or L-SME	methyl <i>L</i> -serinate or <i>L</i> -Serine methyl ester
25	NAC	acetyl- <i>L</i> -cysteine
26	NACME	methyl acetyl- <i>L</i> -cysteinate
27-L	L-pen	(<i>R</i>)-2-amino-3-mercapto-3-methylbutanoic acid
27-D	D-pen	(<i>S</i>)-2-amino-3-mercapto-3-methylbutanoic acid
28	DiD-pen	(2 <i>S</i> ,2' <i>S</i>)-3,3'-disulfanediylbis(2-amino-3-methylbutanoic acid)
29	DiNAC	<i>S</i> -(((<i>R</i>)-2-acetamido-2-carboxyethyl)thio)- <i>N</i> -acetyl- <i>D</i> -cysteine
30	Thiola	(2-mercaptopropanoyl)glycine

Synthesis of Molecular Imposters

Synthesis of L-CDEE (3) and L-CDIE (4). L-cys-Boc (2.0 g) and DMAP (0.1 g) were dissolved in CH_2Cl_2 (20 mL) and ethanol (1.5 mL) in a round bottom flask and stirred for 10 min. DCC (2.8 g) was added to the reaction flask over a period of 30 min. The reaction was stirred overnight and allowed to warm slowly to room temperature. The filtrate was collected after vacuum filtration and then concentrated *in vacuo*. The residue was purified by column chromatography on silica gel (hexane/ethyl acetate = 7:3) to yield L-CDEE-Boc. ^1H NMR (400MHz, CDCl_3): δ = 4.56 (m, 1H), 4.20 (br, 2H), 3.15 (m, 2H) 1.45 (s, 9H), 1.20 (s, 3H). L-CDEE-Boc (1.5 g) was then dissolved in HCl/ethanol (70 mL, 1.25 M) and stirred overnight. The crude mixture was extracted with H_2O (2×200 mL). L-CDEE was obtained as the white hydrochloride salt after lyophilization of the aqueous layer from extraction. L-CDEE was recrystallized from ethanol/ethyl acetate solution and dried under vacuum prior to use: ^1H NMR (400MHz, D_2O): δ = 4.38 (m, 1H), 4.23 (m, 2H), 3.33 (m, 2H), 1.27 (s, 3H). (Figure S1A).

The L-CDIE salt was synthesized according to the same protocol as L-CDEE, using isopropanol instead of ethanol. L-CDIE-Boc: ^1H NMR (400MHz, CDCl_3): δ = 5.07 (m, 1H), 4.56 (m, 1H), 3.17 (m, 2H), 1.45 (s, 9H), 1.16 (d, 6H). L-CDIE salt: ^1H NMR (400MHz, D_2O): δ = 5.08 (m, 1H), 4.45 (m, 1H), 3.30 (m, 2H), 1.25 (d, 6H). (Figure S1B)

Synthesis of L-CDTE (5) perchlorate. 100 mL tert-butyl acetate was added slowly to a solution of L-cystine (10.0 g) dissolved in perchloric acid (70 % aq. v/v, 16.6 mL). The mixture was cooled in an ice bath for 30 min and stirred overnight. The solid was collected via filtration and washed with cold diethyl ether three times, yielding the pure L-CDTE salt: ^1H NMR (400MHz, MeOD_4): δ = 3.68 (br, 1H), 3.12 (m, 1H), 2.88 (m, 1H), 1.49 (s, 9H). (Figure S1C)

Synthesis of L-C=CcysDME (16). 2 mL of thionyl chloride was added to a solution of L-allylglycine salt (1.85 g) dissolved in CH_3OH (24 mL). The mixture was stirred overnight and allowed to slowly warm to room temperature. The solvent was removed *in vacuo* to produce the solid product L-allylglycine methyl ester (L-AGME). ^1H NMR (400MHz, CDCl_3): δ = 5.87 (m, 1H), 5.31 (m, 2H), 4.25 (br, 1H), 3.81 (s, 1H), 2.86 (br, 2H).

Et_3N (2.7 g) was added slowly to a solution of L-AGME (2 g) in CH_2Cl_2 (20 mL) at 0°C with stirring. After stirring for 10 min, di-tert-butyl dicarbonate (T-Boc, 3.6 g) dissolved in CH_2Cl_2 (20 mL) was added dropwise over a period of 30 min. The mixture was stirred overnight and allowed to warm slowly to room temperature, then extracted twice with CH_2Cl_2 (2×200 mL). The combined organic layers were dried over MgSO_4 and concentrated *in vacuo*. The product L-allylglycine-Boc methyl ester (L-AGME-Boc) was

purified by silicon column chromatography (hexane/ethyl acetate = 8:2): ^1H NMR (400MHz, CDCl_3): δ = 5.70 (m, 1H), 5.15 (m, 2H), 5.05 (br, 1H), 4.38 (m, 1H), 3.72 (s, 3H), 2.48 (m, 2H), 1.44 (s, 9H).

L-AGME-Boc (200mg) was dissolved in CH_2Cl_2 (6 mL), and the flask sealed and purged with N_2 . A freshly prepared solution of Grubbs II catalyst (58 mg) dissolved in CH_2Cl_2 (1 mL) was added by syringe and the mixture heated under reflux overnight, after which the solution was concentrated *in vacuo*. The residue was purified by column chromatography on silica gel (hexane/ethyl acetate = 8/2) to yield the compound L-AGME-Boc dimer, (L-C=CcysDME-Boc): ^1H NMR (400MHz, CDCl_3): δ = 5.54(m, 1H), 4.35 (m, 1H), 3.75 (s, 3H), 2.47 (m, 2H), 1.45(s, 18H). L-C=CcysDME-Boc (1.5g) was then dissolved in HCl/ethyl acetate (20 mL, 1.25 M) and stirred overnight. The crude mixture was extracted twice with H_2O (2×200 mL), and L-C=CcysDME was obtained as the white hydrochloride salt after lyophilization of the aqueous layer from extraction. ^1H NMR (400MHz, D_2O): δ = 5.70(m, 1H), 4.20 (m, 1H), 3.87 (s, 3H), 2.72 (m, 2H) (Figure S2A).

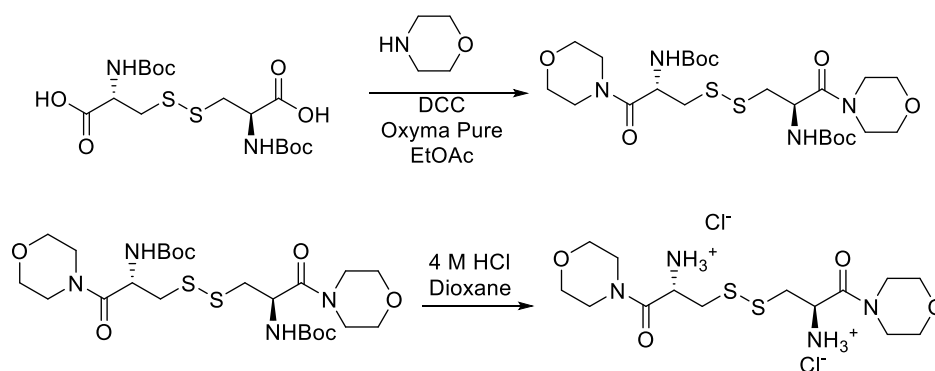
Synthesis of L-CCcysDME (15). L-C=CcysDME-Boc (80 mg) was dissolved in ethyl acetate (20 mL) and 10% Pd/C (40 mg) to the solution added. The reaction flask was then sealed, purged with hydrogen three times, and then shaken overnight under hydrogen (ca. 1 atm). The resulting product was purified using a short column of silica gel to remove the Pd/C catalyst. The residue was concentrated *in vacuo* to yield L-CCcysDME-Boc: ^1H NMR (400MHz, CDCl_3): δ = 5.13(br, 1H), 4.27 (m, 1H), 3.73 (s, 3H), 1.76 (m, 1H), 1.61 (m, 1H), 1.43 (s, 18H), 1.36 (m, 2H). L-CCcysDME-Boc (1.5 g) was then dissolved in HCl/ethyl acetate (20 mL, 1.25M) and stirred overnight, after which the crude mixture was extracted twice with H_2O (2×200 mL). L-CCcysDME was obtained as the white hydrochloride salt after lyophilization of the aqueous layer from extraction and recrystallized from methanol/ethyl acetate solution and dried under vacuum prior to use. ^1H NMR (400MHz, CD_3OD): δ = 3.98 (m, 1H), 3.76 (s, 3H), 1.85 (m, 2H), 1.42 (m, 2H). (Figure S2B)

Synthesis of L-CCcys (14). L-CCcysDME-Boc (0.40 g) was dissolved in CH_3OH (10 mL), followed by the addition of a NaOH solution (0.12 g in 20 mL H_2O). The reaction flask was heated under reflux overnight. The solvent was removed *in vacuo* and ion exchange was performed using 36 (wet) ion-exchange resins for 30 min. L-CCacid-Boc (1.5 g) was dissolved in HCl/ethyl acetate (20 mL, 1.25 M) and stirred overnight. The mixture was extracted twice with H_2O (2×200 mL) and L-CCcys was obtained as the hydrochloride salt after lyophilization. ^1H NMR (400MHz, CD_3OD): δ = 3.98 (m, 1H), 1.97 (m, 2H), 1.56 (m, 2H). (Figure S2C)

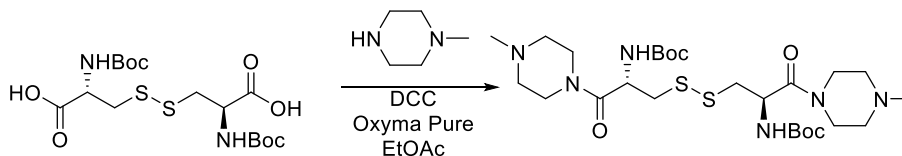
Synthesis of L-HOMOcysDME (18). L-homocystine (0.5 g) was dissolved in 1,4-dioxane (10 mL) and H_2O (20 mL) at 0 °C, followed by the slow addition of a NaOH solution (0.30 g, 7.44 mmol). After stirring for 10 min, di-tert-butyl dicarbonate (T-Boc, 1.02 g) in 1, 4-dioxane (10 mL) was added dropwise

over two hours. The reaction was stirred overnight and allowed to warm slowly to room temperature. The solvent was removed *in vacuo* and 10 mL, HCl (1 M) was added to reduce the pH from pH = 10 to pH = 1, and the mixture was extracted twice with diethyl ether (2×200 mL). The combined organic layers were dried over Na_2SO_4 and the solvent removed *in vacuo* to produce L-homocystine-Boc as a white solid (0.80g). L-homocystine-Boc (0.80 g) and DMAP (0.10 g) were dissolved in CH_2Cl_2 (20 mL), and methanol (2 mL) was added to the mixture, which was then stirred for 10 min. DCC (1.05g) was added to the mixture over 30 min, after which the mixture was allowed to warm slowly to room temperature and then stirred overnight. The filtrate was obtained via filtration and the solvent removed *in vacuo*. The resulting residue was purified by column chromatography on silica gel (hexane/ethyl acetate = 7:3) to yield the compound L-HOMOcysDME-Boc: ^1H NMR (400MHz, CDCl_3): δ = 5.13 (m, 1H), 4.40 (m, 1H), 3.77 (s, 3H), 2.70 (m, 2H), 2.26 (m, 1H), 2.00 (m, 1H), 1.45 (s, 9H). L-HOMOcysDME-Boc (0.5 g) was then dissolved in HCl/ethyl acetate (20 mL, 1.25 M) and stirred overnight. The mixture was extracted with H_2O (2×200 mL), and L-HOMOcysDME was obtained as the hydrochloride salt after lyophilization of the aqueous layer from the extraction. ^1H NMR (400MHz, CD_3OD): δ = 4.25(m, 1H), 3.88(s, 3H), 2.92(m, 2H), 2.39(m, 2H). (Figure S2D)

Synthesis of L-CDMOR (9). The synthesis of compound L-CDMOR (9) was reported in Patent Application US2014/0187546, but is described in more detail here. L-cys-Boc (1.00 g, 2.20 mmol) was added to a 100 ml round bottom flask. Ethyl acetate (70 mL) was added, but the solid did not dissolve. Upon addition of Oxyma (0.71 g, 5.0 mmol) and DCC (1.03 g, 5.0 mmol), the solid dissolved partially and the mixture became yellow. Morpholine (0.90 mL, 9.0 mmol) was then added and the mixture was stirred under argon for up to 24 hours. The mixture was then filtered and the filtrate was washed three times with water (3 x 75 mL). The organic layer was then dried with MgSO_4 and filtered, and the solvent was removed from the filtrate *in vacuo*, affording a viscous oil. The crude oil was purified by silica gel chromatography with ethyl acetate as the eluent, affording the Boc-protected L-CDMOR as a colorless foam (yield = 35 - 50%): ^1H NMR (600 MHz, CDCl_3 , in ppm): δ = 5.46 (d, 2H, 8 Hz), 4.90-4.94 (m, 2H), 3.58-3.74 (m, 16 H), 3.01 (d, 4H, 7 Hz), 1.43 (s, 18H). The Boc-protected L-CDMOR (0.4770 g, 0.81 mmol) and 1 mL of 4 M HCl in dioxane was added to a 25 mL round bottom flask. After stirring for twenty minutes a precipitate was observed and the reaction was stirred for an additional 40 min. The solvent was removed *in vacuo* to afford the hydrochloride salt as a white solid, which was very hygroscopic and formed an oil if left open to the air. ^1H NMR (500 MHz, D_2O) 4.99 (t, 2H, J=6 Hz), 3.70-3.92 (m, 16 H) 3.42 (m, 4 H). ^{13}C NMR (125 MHz, D_2O): 125 MHz 166.1, 66.0, 46.0, 42.8, 37.5.



Synthesis of L-CDNMP (10). The synthesis of compound L-CDNMP (10) using the procedure described in Patent Application US2014/0187546 could not be reproduced with fidelity. Therefore, its synthesis is described here. L-cys-Boc (1.00 g, 2.20 mmol) was added to a 100 ml round bottom flask. 70 mL of ethyl acetate was added, but the solid did not dissolve. Ethyl acetate (70 mL) was added, but the solid did not dissolve. Upon addition of Oxyma (0.71 g, 5.0 mmol) and DCC (1.03 g, 5.0 mmol), the solid dissolved partially and the mixture became yellow. N-methylpiperazine (1.0 mL, 9.0 mmol) was then added and the mixture was stirred under argon for up to 24 hours. The reaction was then filtered and the filtrate was washed three times with water (3 x 75 mL). The organic layer was then dried with MgSO_4 and filtered, and the solvent was removed from the filtrate *in vacuo*, affording a viscous oil. The crude material was purified by silica gel chromatography using a 9:1 ratio of dichloromethane and 10% NH_4OH in methanol. This procedure removed the DCC and DCU, but the product eluted only when 4:1 dichloromethane:10% NH_4OH in methanol was used. The product elutes slowly, requiring collection of numerous fractions after switching to this more polar eluent. The solvent was removed *in vacuo* to afford the Boc-protected L-CDNMP as a white solid, with yields ranging from 25 to 40%. ^1H NMR (400 MHz, CDCl_3 , in ppm): δ = 5.46 (d, 2H, J = 8 Hz), 4.93 (m, 2H), 3.65 (m, 8H), 2.98-3.03 (m, 4H), 2.40-2.48 (m, 8H), 2.32 (s, 6H) 1.43 (s, 18H). ^{13}C NMR (150 MHz, CDCl_3): 168.7, 154.9, 79.9, 54.9, 54.3, 48.8, 45.7, 42.0, 28.2. The Boc-protected L-CDNMP (0.410 g, 0.66 mmol) and 10 mL methanol was added to a 25 mL round bottom flask to afford a colorless solution. Concentrated hydrochloric acid (1 mL) was then added, and the mixture was stirred in air for one hour. The solvent was removed *in vacuo* to afford a white oily solid.



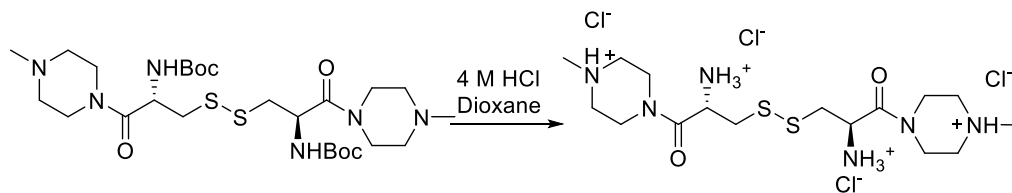


Table S2. Threshold additive concentration, c_{ad} , for the exclusive formation of the L-cystine tetragonal polymorph. The concentration of L-cystine is $c = 3$ mM.

Inhibitor	c_{ad} , at which only hexagonal form crystallizes (mM)	c_{ad} , at which either only tetragonal or mixtures (hexagonal + tetragonal) forms crystallize (mM)	c_{ad} at which only tetragonal always forms ($c_{ad(tetr)}$, mM)
L-HCME	≤ 0.008 (ht7)**	Not observed	≥ 0.010 (t3)
L-CDME	≤ 0.004 (ht22)	0.005 (t1, ht2)	≥ 0.007 (t8)
L-CDPE	≤ 0.0025 (ht6)	0.003 – 0.0035 (t1, ht1)	≥ 0.004 (t1)
L-CDMOR	≤ 0.0005 (ht4)	0.0008 – 0.0015 (t2, ht4)	≥ 0.002 (t6)
L-CDMNP*	≤ 0.00006 (ht5)	0.00008 - 0.0013 (t6, ht7)	≥ 0.0015 (t5)

*4 of 18 experiments with CDMNP were performed at $c = 2.5$ mM.

**Numbers in parentheses indicate the number of experiments for a given c_{ad} range, in which either only tetragonal (t) or hexagonal + tetragonal forms (ht) crystallize.

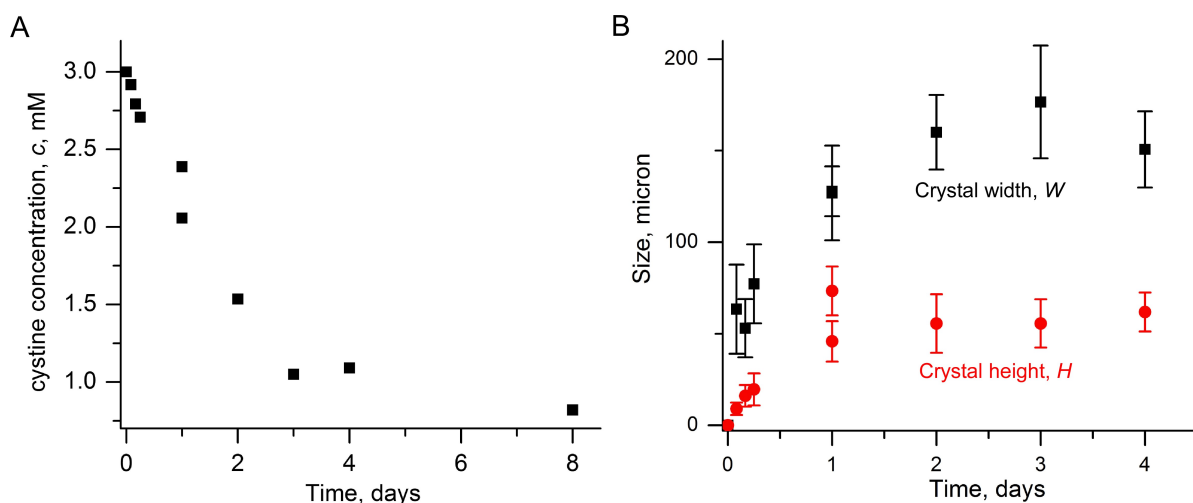


Figure S1. Time evolution of L-cystine concentration (A) and crystal size (B) during bulk crystallization of L-cystine (initial concentration $c = 3$ mM) growth in the presence of $c_{ad} = 0.002$ mM L-CDME.

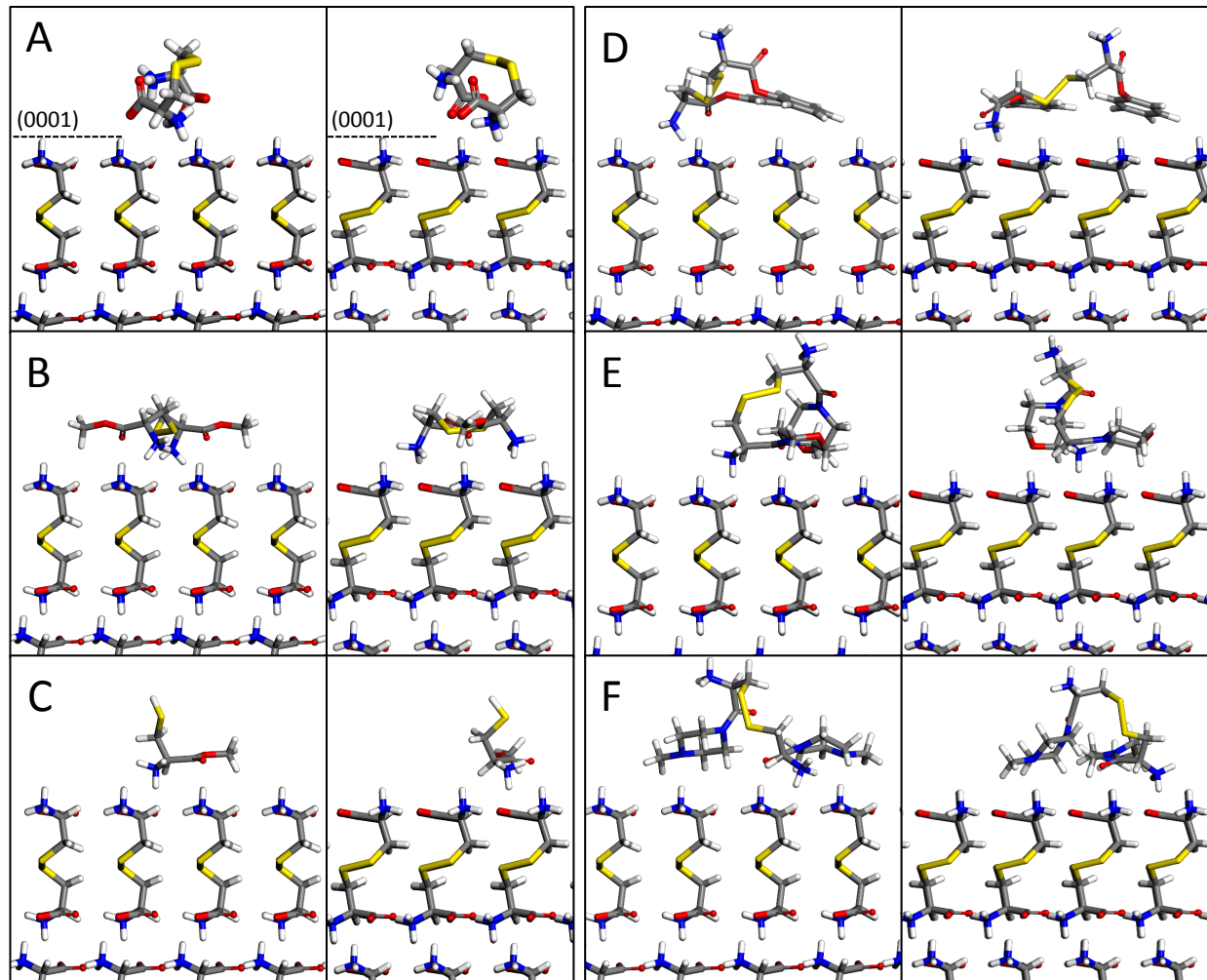


Figure S2. Energetically preferred orientation of adsorbed imposters on the (0001) face. (A) L-cystine, (B) CDME, (C) HCME, (D) CDPE, (E) CDMOR, (F) CDMNP.

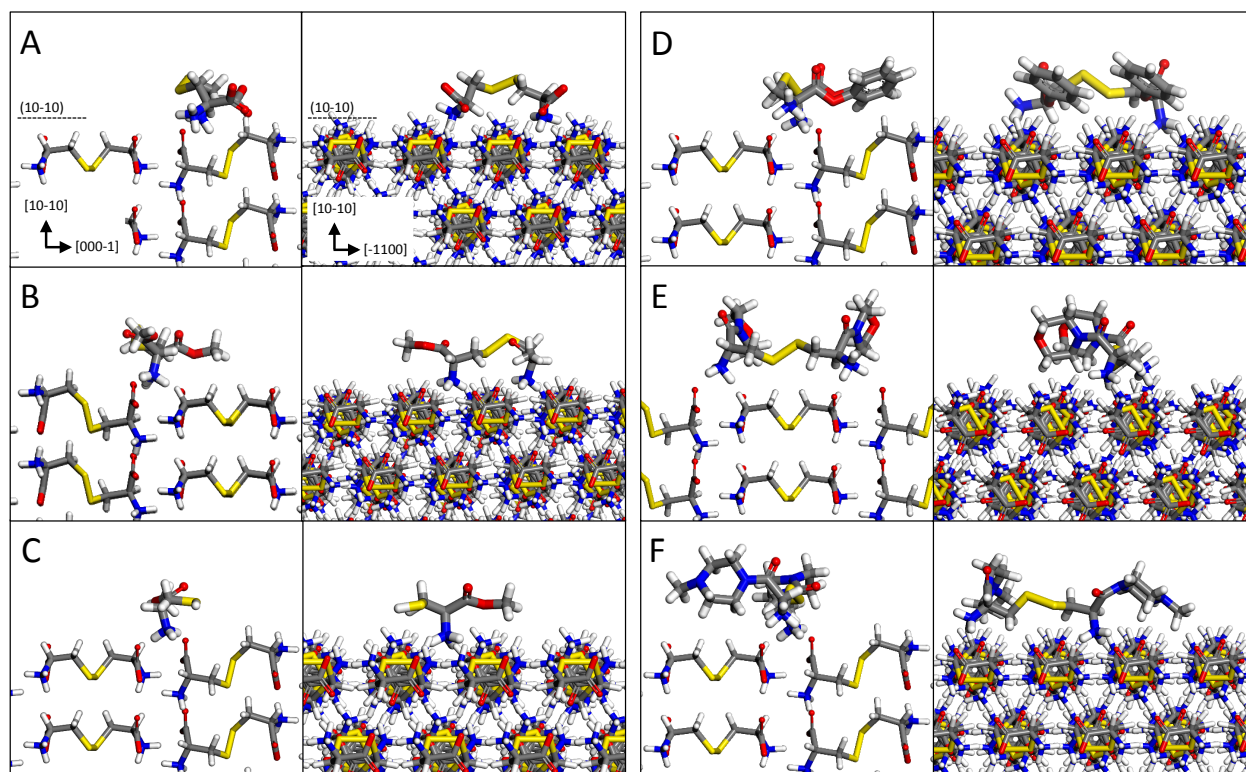


Figure S3. Energetically preferred orientation of adsorbed imposters on the (10-10) face. (A) L-cystine, (B) L-CDME, (C) L-HCME, (D) L-CDPE, (E) L-CDMOR, (F) L-CDNMP.

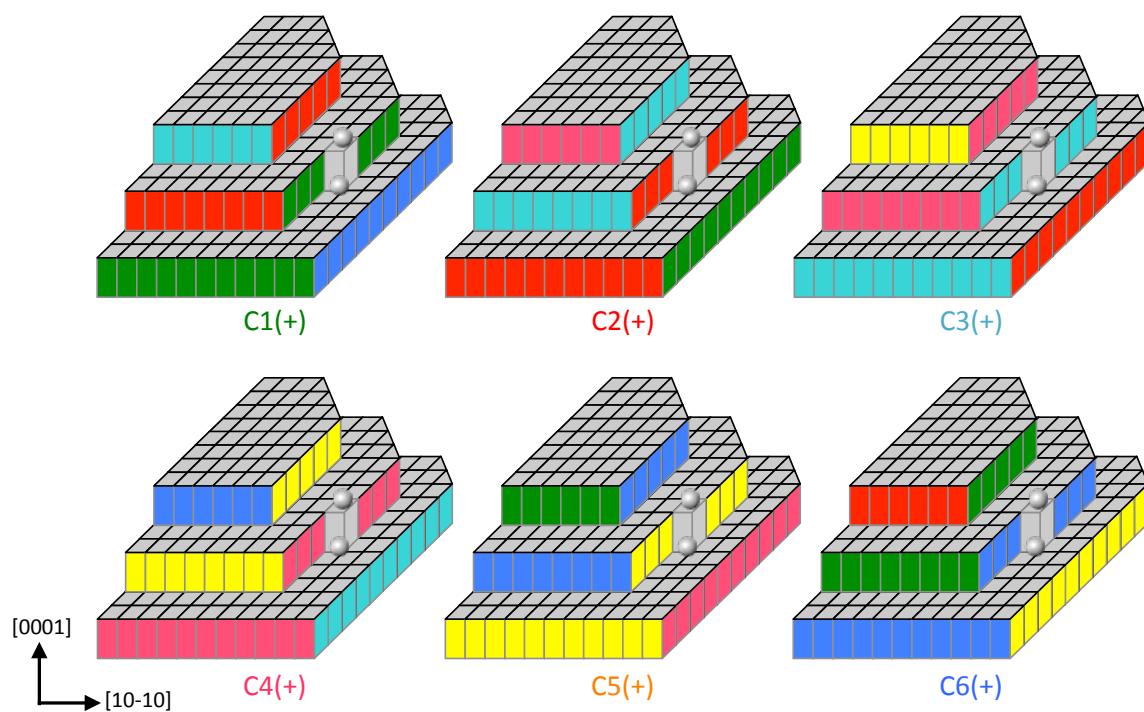


Figure S4. The six unique “clockwise” or “+” kink sites on $\{10\text{-}10\}$ steps on the $\{0001\}$ L-cystine surface. “Counterclockwise” or “-” kink sites not shown.

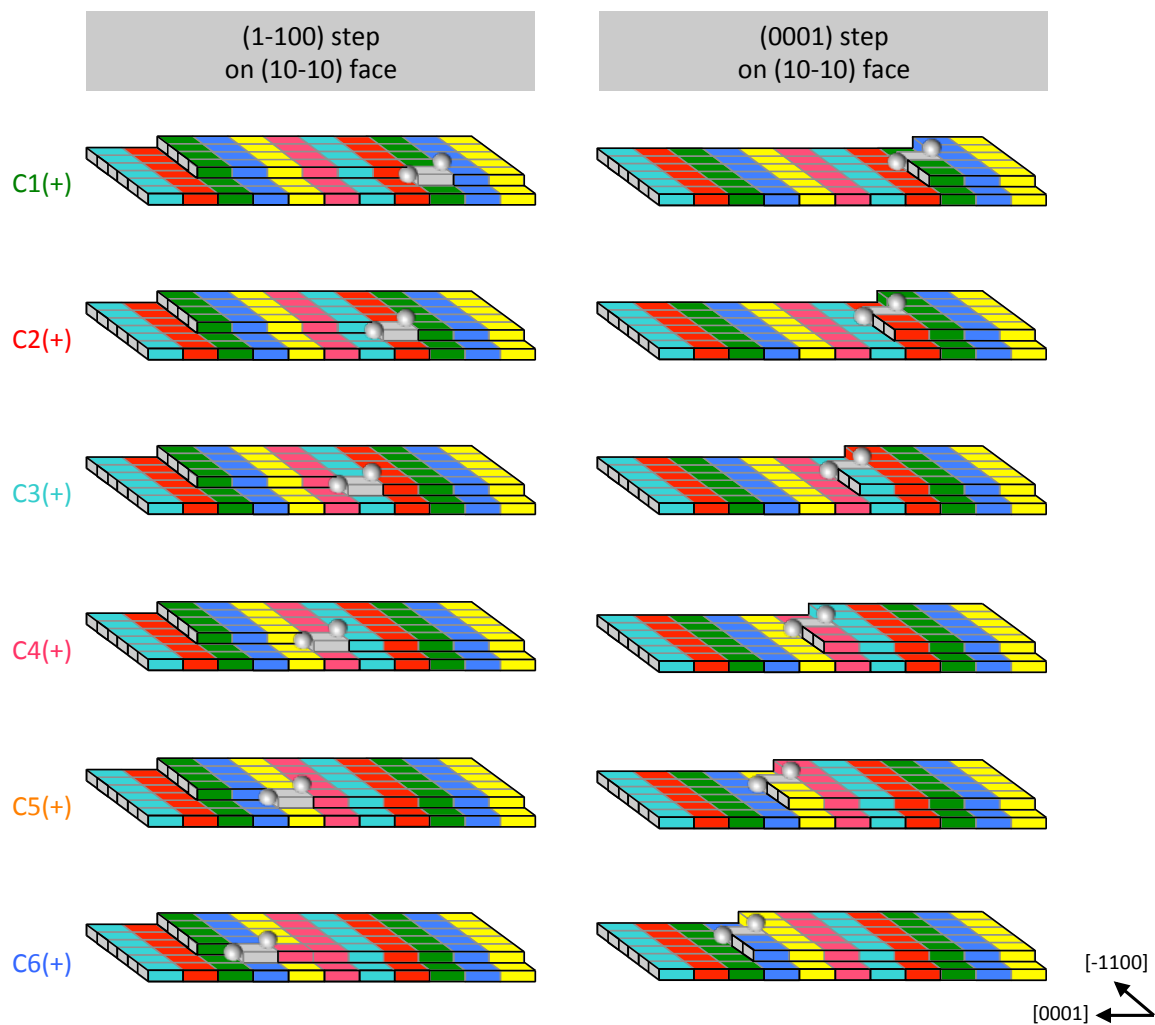


Figure S5. The six unique “clockwise” or “+” kink sites on $\{1-100\}$ steps, on the $\{10-10\}$ L-cystine surface (left), and the six unique “clockwise” or “+” kink sites on the $\{0001\}$ steps on the $\{10-10\}$ L-cystine surface (right). “Counterclockwise” or “-” kink sites not shown.