The chemical ligand space of cereblon

Iuliia Boichenko[1†], Kerstin Bär[1], Silvia Deiss[1], Christopher Heim[1], Reinhard Albrecht[1], Andrei N. Lupas[1], Birte Hernandez Alvarez[1,*], and Marcus D. Hartmann[1,*]

[1] Department of Protein Evolution, Max Planck Institute for Developmental Biology, Max-Planck-Ring 5, 72076 Tübingen, Germany

[†]Current address: Laboratory of Biophysical Chemistry of Macromolecules, Ecole polytechnique fédérale de Lausanne, CH-1015 Lausanne, Switzerland

* Corresponding authors

E-mail: birte.hernandez@tuebingen.mpg.de (BHA) or marcus.hartmann@tuebingen.mpg.de (MDH)

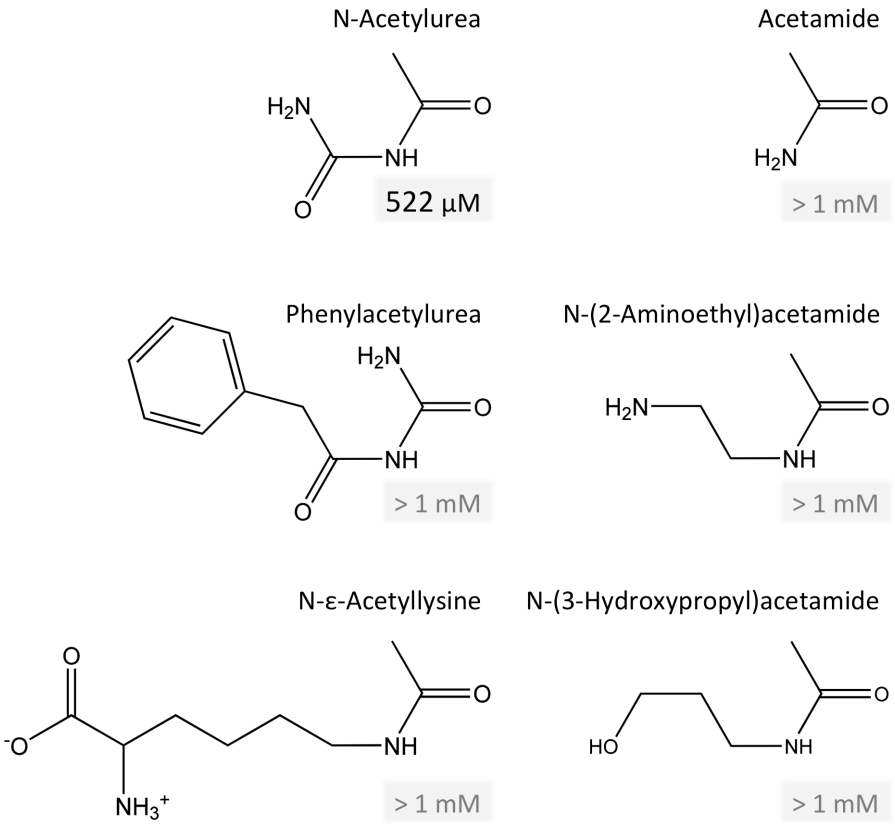


Figure S1

Non-cyclic compounds. Of all tested linear compounds, only acetylurea showed detectable binding to MsCl4.

```
      MsCI4    1 MPLDA---------GGQ---NSTQ  12
CRBN_HUMAN  288 LPIDDVLRIQLLKIGSAIQRLRCE 311

      MsCI4   13 MVLAPGASIFRCRQCGQTISR---  33
CRBN_HUMAN  312 LDIMNKCTSLCCKQCQETEITTKN 335

      MsCI4   34 RDWLLPMGGDHEHVVFNPAGMIFR  57
CRBN_HUMAN  336 EIFSLSLCGP-MAAYVNPHGYVHE 358

      MsCI4   58 VWCFSLAQGLRLIGAPSGEFSWFK  81
CRBN_HUMAN  359 TLTVYKACNLNLIGRPSTEHSWFP 382

      MsCI4   82 GYDWTIALCGQCGSHLGWHYEGG- 104
CRBN_HUMAN  383 GYAWTVAQCKICASHIGWKFTATK 406

      MsCI4  105 --SQPQTFFGLIKDRLAEGPAD   124
CRBN_HUMAN  407 KDMSPQKFWGLTRSALLPTIPD   428
```

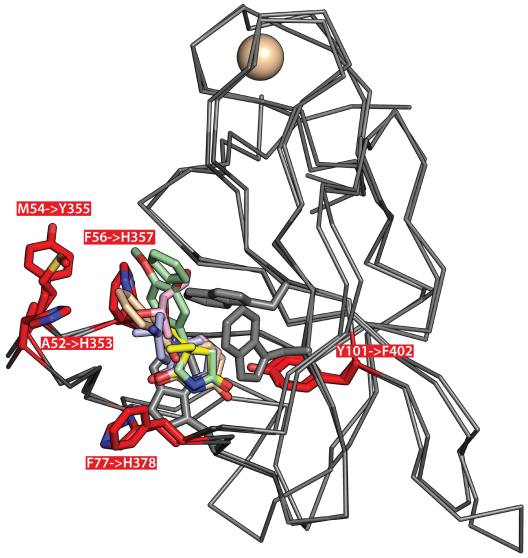
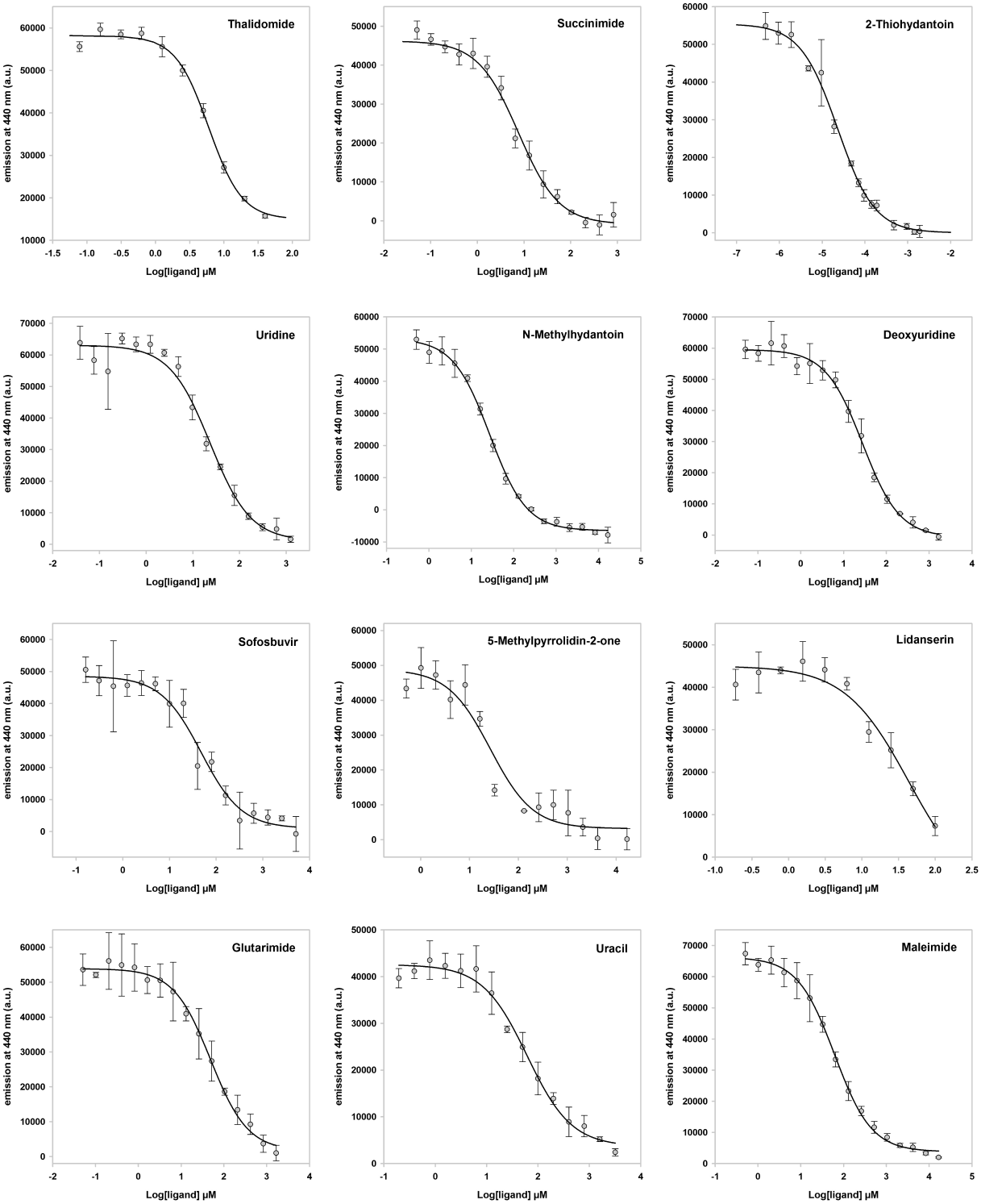


Figure S2

Sequence alignment and structural superposition of human cereblon and MsCI4 in complex with the effectors ethosuximide (yellow), deoxyuridine (pink), aminoglutethimide (blue), thalidomide (brown) and rolipram (green), highlighting species-specific differences in the vicinity of the thalidomide-binding site. Non-conserved side chains within a 4Å radius around all atoms of all effectors are highlighted red. From the highlighted 5 substitutions, Y101 and F77 affect only the binding moiety and thus have the same effect on all effectors. The other three residues are in the 4Å sphere of the protruding moieties but do not interact with any of them.


Thalidomide
Succinimide
2-Thiohydantoin
Uridine
N-Methylhydantoin
Deoxyuridine
Sofosbuvir
5-Methylpyrrolidin-2-one
Lidanserin
Glutarimide
Uracil
Maleimide
emission at 440 nm (a.u.)
Log[ligand] µM

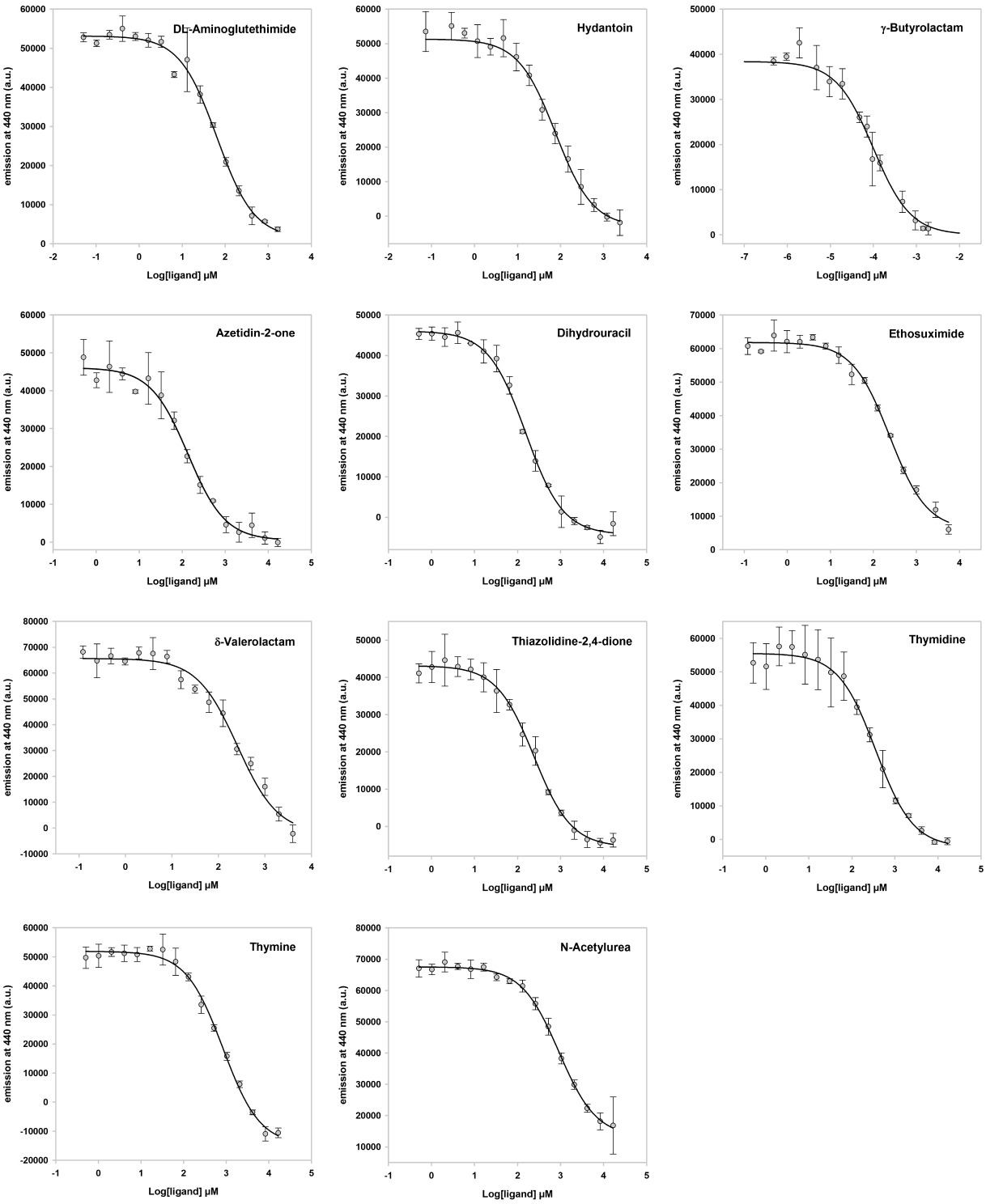


Figure S3

FRET curves for all identified binders.

Substance	Ki [µM]
Succinimide (Pyrrolidine-2,5-dione)	4.3 ± 0.5
(±)-Thalidomide	4.4 ± 0.2
2-Thiohydantoin	13 ± 1.3
Uridine	13 ± 1.9
Deoxyuridine	15 ± 1.6
N-Methylhydantoin	16 ± 1.3
5-Methylpyrrolidin-2-one	19 ± 6.0
Sofosbuvir	22 ± 4.4
Lidanserin	25 ± 8.5
Glutarimide (Piperidine-2,6-dione)	28 ± 5.0
Uracil (Pyrimidine-2,4-dione)	35 ± 5.1
Maleimide (Pyrrol-2,5-dione)	35 ± 3.1
DL-Aminoglutethimide	39 ± 3.8
Hydantoin (Imidazolidine-2,4-dione)	43 ± 10.7
γ-Butyrolactam (Pyrrolidin-2-one)	57 ± 11.6
Azetidin-2-one	77 ± 11.0
Dihydrouracil (Hexahydropyrimidine-2,4-dione)	88 ± 6.8
Ethosuximide (3-Ethyl-3-methylsuccinimide)	136 ± 11.6
δ-Valerolactam (Piperidin-2-one)	146 ± 20.3
Thiazolidine-2,4-dione	137 ± 15.9
Thymidine	204 ± 34.5
Thymine (5-Methylpyrimidine-2,4-dione)	478 ± 41.6
N-Acetylurea	522 ± 53.5
Acetamide	> 1,000
1-Azo-2-Cyclooctanone	> 1,000
Barbiturate	> 1,000
Creatinine (2-Imino-N-methylhydantoin)	> 1,000
Cycloheximide	> 1,000
Cytosine	> 1,000
5,5-Dimethyl-oxazolidine-2,4-dione	> 1,000
ε-Caprolactam	> 1,000

5-Ethyl-5-methyl-hydantoin	> 1,000
γ-Butyrolactone	> 1,000
Glutaric Anhydride	> 1,000
N-(2-Aminoethyl)acetamide	> 1,000
N-ε-Acetyllysine	> 1,000
N-(2-Hydroxypropyl)acetamide	> 1,000
Phenylacetylurea	> 1,000
Piperidine	> 1,000
Pseudouridine	> 1,000
Pyrrolidine	> 1,000
Pyrrolidine-2-thione	> 1,000
Succinic Anhydride	> 1,000
Dasabuvir	n.d.
Dantrolene	n.d.
Dexetimide	n.d.
Glutethimide	unavailable
Nitrofurantoin	n.d.
Oxazolidine-2,4-dione	unavailable
Rogletimide	unavailable
Rolipram	n.d. – X-tal

Table S1

All substances addressed in this study, sorted by affinity for MsCl4.

n.d.: not determined due to solubility issues

X-tal: binding verified via crystallography

Ligand	Amino-glutethimide	γ-Butyro-lactam	Etho-suximide	Glutarimide	Hydantoin	Rolipram	Thiazolidine-2,4-dione	2-Thio-hydantoin	δ-Valerolactam
Data collection									
Cell dimensions *a*, *b*, *c* (Å)	56.8, 60.4, 88.5	56.9, 59.7 88.4	56.9, 59.8, 88.1	56.2, 59.7, 88.3	56.4, 58.5, 87.1	57.3, 59.5, 88.0	56.5, 59.9, 88.6	56.6, 59.9, 88.3	56.5, 59.6, 88.6
Resolution (Å)	37.5 - 1.70 (1.80 – 1.70)	37.4 – 1.90 (2.01 – 1.90)	37.3 – 2.10 (2.22 – 2.10)	37.1 – 2.30 (2.44 – 2.30)	36.8 – 1.85 (1.96 – 1.85)	37.4 – 1.95 (2.07 – 1.95)	37.3 – 1.65 (1.75 – 1.65)	37.3 – 1.55 (1.64 – 1.55)	37.2 – 1.70 (1.80 – 1.70)
R_{merge}	4.3 (59.2)	5.4 (48.7)	6.5 (48.6)	8.0 (36.9)	6.1 (59.5)	7.6 (65.4)	5.8 (54.7)	4.7 (50.7)	4.7 (49.4)
$I/\sigma I$	15.4 (1.97)	12.5 (1.89)	13.4 (2.29)	8.64 (1.98)	12.6 (1.86)	10.8 (1.91)	10.1 (1.71)	13.8 (2.4)	13.3 (2.11)
Completeness (%)	99.7 (99.2)	98.4 (96.7)	99.2 (97.1)	97.0 (94.0)	97.8 (94.6)	99.8 (99.4)	99.1 (97.5)	99.4 (98.5)	99.2 (97.7)
Redundancy	3.45 (3.42)	2.86 (2.81)	3.39 (3.24)	2.55 (2.52)	2.92 (2.87)	4.64 (4.63)	2.87 (2.68)	2.85 (3.39)	3.23 (3.20)
Refinement									
Resolution (Å)	37.5 - 1.70 (1.74 – 1.70)	37.3 – 1.90 (1.95 – 1.90)	37.3 – 2.10 (2.15 – 2.10)	37.1 – 2.30 (2.36 – 2.30)	35.0 – 1.85 (1.90 – 1.85)	37.4 – 1.95 (2.0 – 1.95)	37.3 – 1.65 (1.69 – 1.65)	37.3 – 1.55 (1.59 – 1.55)	37.2 – 1.70 (1.74 – 1.70)
No. reflections	32512	23182	17207	12940	23726	21463	35109	41911	31914
R_{work}	0.17 (0.26)	0.18 (0.27)	0.18 (0.23)	0.19 (0.29)	0.16 (0.24)	0.18 (0.27)	0.17 (0.29)	0.16 (0.26)	0.17 (0.27)
R_{free}	0.21 (0.27)	0.22 (0.31)	0.23 (0.25)	0.25 (0.35)	0.20 (0.30)	0.22 (0.30)	0.21 (0.26)	0.20 (0.33)	0.21 (0.26)
Ligand in chain(s)	C	A, B, C	A, B	A, B, C	A, B	A, B	A, B, C	A, B, C	A, B, C
PDB code	5OH1	5OH2	5OH3	5OH4	5OH7	5OH8	5OH9	5OHA	5OHB

Table S2

Data collection and refinement statistics. Highest resolution shells are shown in parenthesis.